Katcheves, Konstantina

From: Fredman, Jeffrey
Sent: Thursday, December 15, 2005 12:34 PM
To: STIC-Biotech/ChemLib
Cc: Katcheves, Konstantina
Subject: FW: 10-826573

PLEASE RUSH.

I Approve.

If you do not understand the search, please contact me for an explanation.

Jeff Fredman

-----Original Message-----
From: Katcheves, Konstantina
Sent: Thursday, December 15, 2005 9:47 AM
To: Fredman, Jeffrey
Subject: FW: 10-826573

Jeff:

Christina is out. Could you approve this RUSH search?

Thanks,
Tina

-----Original Message-----
From: Katcheves, Konstantina
Sent: Thursday, December 15, 2005 9:23 AM
To: Chan, Christina
Subject: RE: 10-826573

Christina:

Would you approve the following search to STIC?

I need to search SEQ ID NO:2 with mutations at positions 54, 242, and 372. The mutations can be any one of A, T, G, C. In the email below Jeff suggested a search of the 18mers with each posible mutant at each position. Can you do this? If not what do you suggest?

Thanks,
Tina

-----Original Message-----

From: Fredman, Jeffrey
Sent: Wednesday, December 14, 2005 6:30 AM
To: Katcheves, Konstantina
Subject: RE: 10-826573

I would search the 18 mers or so overlapping those positions with each possible mutant (so if position 54 can be an A or a C, I would request a nnnnnnnnnAnnnnnn and nnnnnnnCnnnnnnn type search). That should pick up full length sequences to either mutation, as well as some oligos (but not necessarily all oligos). Depending on the claims, other searches might also be appropriate.

Jeff

-----Original Message-----
From: Katcheves, Konstantina
Sent: Monday, December 12, 2005 12:27 PM
To: Fredman, Jeffrey
Subject: 10-826573

Jeff:

How would you search a claim to mutants of SEQ ID NO:2 having mutations at positions 54, 242, and 372?

Thanks,
Tina

Konstantina Katcheves
Patent Examiner , AU1636
Phone: (571) 272-0768
Room: REM 2A60
Mail: REM 2C70

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:53 ; Search time 1597.5 Seconds
(without alignments)
452.721 Million cell updates/sec

Title: US-10-826-573-5
Perfect score: 19
Sequence: 1 ctgtctcttatacacatct 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	19	100.0	101	6	CD110771	CD110771 ME1-0021T
c 2	19	100.0	117	6	CD150079	CD150079 ML1-0017T
c 3	19	100.0	156	4	BG217173	BG217173 RST36874
c 4	19	100.0	159	9	CG992997	CG992997 BAC#13_R4
c 5	19	100.0	181	9	CG992974	CG992974 BAC#13_R4
c 6	19	100.0	252	9	CG681064	CG681064 BAC#16_R4
c 7	19	100.0	285	9	CG992975	CG992975 BAC#13_R4
c 8	19	100.0	289	9	CG681054	CG681054 BAC#16_R4
c 9	19	100.0	329	9	CG992989	CG992989 BAC#13_R4
c 10	19	100.0	342	9	CG681053	CG681053 BAC#16_R4
c 11	19	100.0	362	9	CG992990	CG992990 BAC#13_R4
c 12	19	100.0	368	9	CG681071	CG681071 BAC#16_R4
c 13	19	100.0	369	9	CG681056	CG681056 BAC#16_R4
c 14	19	100.0	375	9	CG681072	CG681072 BAC#16_R4
c 15	19	100.0	422	9	CG681048	CG681048 BAC#16_R4
c 16	19	100.0	437	9	CG992988	CG992988 BAC#13_R4
c 17	19	100.0	444	9	CG992981	CG992981 BAC#13_R4
c 18	19	100.0	449	9	CG681074	CG681074 BAC#16_R4
c 19	19	100.0	489	9	CG681055	CG681055 BAC#16_R4
c 20	19	100.0	500	9	CG681070	CG681070 BAC#16_R4
c 21	19	100.0	510	9	CG681057	CG681057 BAC#16_R4
c 22	19	100.0	522	9	CG992971	CG992971 BAC#13_R4
c 23	19	100.0	523	9	CG992972	CG992972 BAC#13_R4
c 24	19	100.0	529	9	CC902784	CC902784 R4-8-BAC#
c 25	19	100.0	529	9	CG992978	CG992978 BAC#13_R4
c 26	19	100.0	535	9	CG992983	CG992983 BAC#13_R4
c 27	19	100.0	546	9	CG681058	CG681058 BAC#16_R4
c 28	19	100.0	553	9	CG992973	CG992973 BAC#13_R4
c 29	19	100.0	573	9	CG992979	CG992979 BAC#13_R4
c 30	19	100.0	575	9	CG992982	CG992982 BAC#13_R4
c 31	19	100.0	589	9	CG993003	CG993003 BAC#13_R4
c 32	19	100.0	593	9	CG993011	CG993011 BAC#13_R4
c 33	19	100.0	598	9	CG992976	CG992976 BAC#13_R4
c 34	19	100.0	598	9	CG992977	CG992977 BAC#13_R4
c 35	19	100.0	612	9	CG681052	CG681052 BAC#16_R4
c 36	19	100.0	618	9	CG993008	CG993008 BAC#13_R4
c 37	19	100.0	620	9	CG681059	CG681059 BAC#16_R4
c 38	19	100.0	620	9	CG992998	CG992998 BAC#13_R4
c 39	19	100.0	630	9	CG992980	CG992980 BAC#13_R4
c 40	19	100.0	640	9	CG993007	CG993007 BAC#13_R4
c 41	19	100.0	653	9	CG992991	CG992991 BAC#13_R4
c 42	19	100.0	661	9	CG993001	CG993001 BAC#13_R4
c 43	19	100.0	667	9	CG992994	CG992994 BAC#13_R4
c 44	19	100.0	667	9	CG993006	CG993006 BAC#13_R4
c 45	19	100.0	675	9	CG992970	CG992970 BAC#13_R4

ALIGNMENTS

RESULT 1
CD110771/c
LOCUS CD110771 101 bp mRNA linear EST 14-SEP-2003
DEFINITION ME1-0021T-D155-C11-U.G ME1-0021 Schistosoma mansoni cDNA clone ME1-0021T-D155-C11.G, mRNA sequence.
ACCESSION CD110771
VERSION CD110771.1 GI:34650114
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 101)
AUTHORS Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M., Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr., Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F., Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L., Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A., Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A., Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T., Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M., Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
TITLE Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni
JOURNAL Nat. Genet. 35 (2), 148-157 (2003)
MEDLINE 22879926
PUBMED 12973350
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: ME1-0021T-D155 row: 11 column: C.
FEATURES Location/Qualifiers
source 1. .101
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="ME1-0021T-D155-C11.G"
/sex="mixed pool"

```
                     /dev_stage="egg"
                     /lab_host="Mus musculus"
                     /clone_lib="ME1-0021"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 101;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         32 CTGTCTCTTATACACATCT 14


RESULT 2
CD150079/c
LOCUS       CD150079                 117 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ML1-0017T-R147-G07-U.G ML1-0017 Schistosoma mansoni cDNA clone
            ML1-0017T-R147-G07.G, mRNA sequence.
ACCESSION   CD150079
VERSION     CD150079.1  GI:34687822
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 117)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
  MEDLINE   22879926
   PUBMED   12973350
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ML1-0017T-R147  row: 7  column: G.
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="ML1-0017T-R147-G07.G"
                     /sex="mixed pool"
                     /dev_stage="miracidium"
                     /clone_lib="ML1-0017"
                     /note="Vector: pGEM T-easy"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 117;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         57 CTGTCTCTTATACACATCT 39


RESULT 3
BG217173/c
LOCUS       BG217173                 156 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST36874 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG217173
VERSION     BG217173.1  GI:13743194
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 156)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
            Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozer,S.,
            Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
            Offenbacher,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
   PUBMED   11329013
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 156.
FEATURES             Location/Qualifiers
     source          1. .156
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HT1080"
                     /clone_lib="Athersys RAGE Library"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 156;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         52 CTGTCTCTTATACACATCT 34


RESULT 4
CG992997/c
LOCUS       CG992997                 159 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_D03r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            D03r, genomic survey sequence.
ACCESSION   CG992997
VERSION     CG992997.1  GI:39946882
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 159)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
```

```
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .159
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="D03r"
                      /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 159;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         42 CTGTCTCTTATACACATCT 24


RESULT 5
CG992974/c
LOCUS       CG992974                181 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A09f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A09f, genomic survey sequence.
ACCESSION   CG992974
VERSION     CG992974.1  GI:39946855
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 181)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .181
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="A09f"
                      /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 181;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 6
CG681064/c
LOCUS       CG681064                252 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_A01 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone A01, genomic survey sequence.
ACCESSION   CG681064
VERSION     CG681064.1  GI:37577901
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 252)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .252
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="A01"
                      /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         26 CTGTCTCTTATACACATCT 8


RESULT 7
CG992975/c
LOCUS       CG992975                285 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A09r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A09r, genomic survey sequence.
ACCESSION   CG992975
VERSION     CG992975.1  GI:39946856
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 285)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .285
```

```
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A09r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         41 CTGTCTCTTATACACATCT 23


RESULT 8
CG681054/c
LOCUS       CG681054                289 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_C06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone C06, genomic survey sequence.
ACCESSION   CG681054
VERSION     CG681054.1  GI:37577891
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 289)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         24 CTGTCTCTTATACACATCT 6


RESULT 9
CG992989/c
LOCUS       CG992989                329 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C08r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C08r, genomic survey sequence.
ACCESSION   CG992989
VERSION     CG992989.1  GI:39946872
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C08r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 329;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         40 CTGTCTCTTATACACATCT 22


RESULT 10
CG681053/c
LOCUS       CG681053                342 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_B06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone B06, genomic survey sequence.
ACCESSION   CG681053
VERSION     CG681053.1  GI:37577890
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN-2R
            Seq primer: 5' GCAATGTAACATCAGAGATTTTGAG 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="B06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 342;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         41 CTGTCTCTTATACACATCT 23


RESULT 11
CG992990/c
LOCUS       CG992990                362 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C09f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C09f, genomic survey sequence.
ACCESSION   CG992990
VERSION     CG992990.1  GI:39946875
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 362)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .362
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C09f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 12
CG681071/c
LOCUS       CG681071                368 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_C02 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone C02, genomic survey sequence.
ACCESSION   CG681071
VERSION     CG681071.1  GI:37577908
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .368
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C02"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 368;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 13
CG681056/c
LOCUS       CG681056                369 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_E06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone E06, genomic survey sequence.
ACCESSION   CG681056
VERSION     CG681056.1  GI:37577893
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .369
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="E06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         27 CTGTCTCTTATACACATCT 9


RESULT 14
CG681072/c
LOCUS       CG681072                375 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_E03 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone E03, genomic survey sequence.
ACCESSION   CG681072
```

```
VERSION     CG681072.1  GI:37577909
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 375)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="E03"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 9;  Length 375;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         25 CTGTCTCTTATACACATCT 7


RESULT 15
CG681048/c
LOCUS       CG681048                422 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_F07 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone F07, genomic survey sequence.
ACCESSION   CG681048
VERSION     CG681048.1  GI:37577885
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="F07"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 9;  Length 422;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         26 CTGTCTCTTATACACATCT 8


RESULT 16
CG992988/c
LOCUS       CG992988                437 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C08f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C08f, genomic survey sequence.
ACCESSION   CG992988
VERSION     CG992988.1  GI:39946871
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 437)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .437
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C08f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 9;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 17
CG992981/c
LOCUS       CG992981                444 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C04f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C04f, genomic survey sequence.
ACCESSION   CG992981
VERSION     CG992981.1  GI:39946864
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
```

```
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .444
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="C04f"
                      /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 444;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 18
CG681074/c
LOCUS       CG681074                449 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_G03 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone G03, genomic survey sequence.
ACCESSION   CG681074
VERSION     CG681074.1  GI:37577911
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .449
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="G03"
                      /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         26 CTGTCTCTTATACACATCT 8


RESULT 19
CG681055/c
LOCUS       CG681055                489 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_D06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone D06, genomic survey sequence.
ACCESSION   CG681055
VERSION     CG681055.1  GI:37577892
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN-2R
            Seq primer: 5' GCAATGTAACATCAGAGATTTTGAG 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .489
                      /organism="Zea mays"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:4577"
                      /clone="D06"
                      /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         36 CTGTCTCTTATACACATCT 18


RESULT 20
CG681070/c
LOCUS       CG681070                500 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_D12 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone D12, genomic survey sequence.
ACCESSION   CG681070
VERSION     CG681070.1  GI:37577907
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 500)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN-2R
            Seq primer: 5' GCAATGTAACATCAGAGATTTTGAG 3'
            Class: transposon-tagged.
FEATURES              Location/Qualifiers
     source           1. .500
```

```
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="D12"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 500;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         35 CTGTCTCTTATACACATCT 17


RESULT 21
CG681057/c
LOCUS       CG681057                  510 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_F06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone F06, genomic survey sequence.
ACCESSION   CG681057
VERSION     CG681057.1  GI:37577894
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 510)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN-2R
            Seq primer: 5' GCAATGTAACATCAGAGATTTTGAG 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .510
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="F06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 510;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         40 CTGTCTCTTATACACATCT 22


RESULT 22
CG992971/c
LOCUS       CG992971                  522 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A07r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A07r, genomic survey sequence.
ACCESSION   CG992971
VERSION     CG992971.1  GI:39946852
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A07r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 522;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         29 CTGTCTCTTATACACATCT 11


RESULT 23
CG992972/c
LOCUS       CG992972                  523 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A08f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A08f, genomic survey sequence.
ACCESSION   CG992972
VERSION     CG992972.1  GI:39946853
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A08f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 523;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 24
CC902784/c
LOCUS       CC902784                529 bp    DNA     linear   GSS 08-AUG-2003
DEFINITION  R4-8-BAC#16-KAN-2F-C03 Zea mays BSH-R4-8-#13/#16-BACs Zea mays
            genomic clone C03, genomic survey sequence.
ACCESSION   CC902784
VERSION     CC902784.1  GI:33521717
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C03"
                     /clone_lib="Zea mays BSH-R4-8-#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         74 CTGTCTCTTATACACATCT 56


RESULT 25
CG992978/c
LOCUS       CG992978                529 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A11r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A11r, genomic survey sequence.
ACCESSION   CG992978
VERSION     CG992978.1  GI:39946859
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A11r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         38 CTGTCTCTTATACACATCT 20


RESULT 26
CG992983/c
LOCUS       CG992983                535 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C05f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C05f, genomic survey sequence.
ACCESSION   CG992983
VERSION     CG992983.1  GI:39946866
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C05f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 535;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 27
CG681058/c
LOCUS       CG681058                546 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_H06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone H06, genomic survey sequence.
ACCESSION   CG681058
```

```
VERSION     CG681058.1  GI:37577895
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN-2R
            Seq primer: 5' GCAATGTAACATCAGAGATTTTGAG 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="H06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 546;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         39 CTGTCTCTTATACACATCT 21


RESULT 28
CG992973/c
LOCUS       CG992973                 553 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A08r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A08r, genomic survey sequence.
ACCESSION   CG992973
VERSION     CG992973.1  GI:39946854
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A08r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         38 CTGTCTCTTATACACATCT 20


RESULT 29
CG992979/c
LOCUS       CG992979                 573 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A12f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A12f, genomic survey sequence.
ACCESSION   CG992979
VERSION     CG992979.1  GI:39946860
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A12f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 573;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 30
CG992982/c
LOCUS       CG992982                 575 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_C04r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            C04r, genomic survey sequence.
ACCESSION   CG992982
VERSION     CG992982.1  GI:39946865
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 575)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
```

```
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .575
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="C04r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 575;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         32 CTGTCTCTTATACACATCT 14


RESULT 31
CG993003/c
LOCUS       CG993003                589 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_D07r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            D07r, genomic survey sequence.
ACCESSION   CG993003
VERSION     CG993003.1  GI:39946890
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="D07r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 589;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         38 CTGTCTCTTATACACATCT 20


RESULT 32
CG993011/c
LOCUS       CG993011                593 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_E02f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            E02f, genomic survey sequence.
ACCESSION   CG993011
VERSION     CG993011.1  GI:39946900
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .593
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="E02f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 593;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         31 CTGTCTCTTATACACATCT 13


RESULT 33
CG992976/c
LOCUS       CG992976                598 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A10f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A10r, genomic survey sequence.
ACCESSION   CG992976
VERSION     CG992976.1  GI:39946857
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
```

```
     source          1. .598
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A10r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 34
CG992977/c
LOCUS       CG992977                598 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A11f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A11f, genomic survey sequence.
ACCESSION   CG992977
VERSION     CG992977.1  GI:39946858
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .598
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A11f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 598;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 35
CG681052/c
LOCUS       CG681052                612 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_A06 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone A06, genomic survey sequence.
ACCESSION   CG681052
VERSION     CG681052.1  GI:37577889
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .612
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A06"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 612;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         26 CTGTCTCTTATACACATCT 8


RESULT 36
CG993008/c
LOCUS       CG993008                618 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_D12f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            D12f, genomic survey sequence.
ACCESSION   CG993008
VERSION     CG993008.1  GI:39946895
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="D12f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 618;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         28 CTGTCTCTTATACACATCT 10


RESULT 37
CG681059/c
LOCUS       CG681059                620 bp    DNA     linear   GSS 08-OCT-2003
DEFINITION  BAC#16_R4-8_A07 Zea mays BSH_Ring4-8_#13/#16-BACs Zea mays genomic
            clone A07, genomic survey sequence.
ACCESSION   CG681059
VERSION     CG681059.1  GI:37577896
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN-2F
            Seq primer: 5' ACCTACAACAAAGCTCTCATCAACC 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A07"
                     /clone_lib="Zea mays BSH_Ring4-8_#13/#16-BACs"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 620;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         20 CTGTCTCTTATACACATCT 2


RESULT 38
CG992998/c
LOCUS       CG992998                620 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_D04f Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            D04f, genomic survey sequence.
ACCESSION   CG992998
VERSION     CG992998.1  GI:39946885
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            FORWARD: KAN2f
            Seq primer: 5' - ACCTACAACAAAGCTCTCATCAACC - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="D04f"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 620;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         27 CTGTCTCTTATACACATCT 9


RESULT 39
CG992980/c
LOCUS       CG992980                630 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_A12r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            A12r, genomic survey sequence.
ACCESSION   CG992980
VERSION     CG992980.1  GI:39946861
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="A12r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         30 CTGTCTCTTATACACATCT 12


RESULT 40
CG993007/c
LOCUS       CG993007                640 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  BAC#13_R4-8_D10r Zea mays BSH_Ring4-8_BAC#13 Zea mays genomic clone
            D10r, genomic survey sequence.
```

```
ACCESSION   CG993007
VERSION     CG993007.1  GI:39946894
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Phelps,T.L., Theuri,J.M. and Birchler,J.A.
  TITLE     Sequence from a B-specific hybridizing BAC
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Birchler, JA
            University of Missouri
            117 Tucker Hall, Columbia, MO 65211, USA
            Tel: 5738824905
            Fax: 5738820123
            Email: birchlerj@missouri.edu
            PCR PRimers
            BACKWARD: KAN2r
            Seq primer: 5' - GCAATGTAACATCAGAGATTTTGAG - 3'
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4577"
                     /clone="D10r"
                     /clone_lib="Zea mays BSH_Ring4-8_BAC#13"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 9;  Length 640;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         25 CTGTCTCTTATACACATCT 7


Search completed: June 13, 2005, 11:36:16
Job time : 1598.5 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:52 ; Search time 200.5 Seconds
(without alignments)
560.973 Million cell updates/sec

Title: US-10-826-573-5
Perfect score: 19
Sequence: 1 ctgtctcttatacacatct 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	Query Match %	Length	DB	ID	Description
	1	19	100.0	19	2	AAV28400	Aav28400 Transposo
	2	19	100.0	19	2	AAZ06436	Aaz06436 Mutant Ou
	3	19	100.0	19	3	AAA11737	Aaa11737 Transposo
	4	19	100.0	19	4	AAD21282	Aad21282 Mosaic te
	5	19	100.0	19	4	AAC91688	Aac91688 Transposo
	6	19	100.0	19	9	ADA13512	Ada13512 Outer end
	7	19	100.0	19	10	AAD58809	Aad58809 Tn5 trans
c	8	19	100.0	19	10	AAD58810	Aad58810 Tn5 trans
	9	19	100.0	19	10	ADF79368	Adf79368 Transposo
	10	19	100.0	19	12	ADM95006	Adm95006 Inverted
c	11	19	100.0	19	12	ADM95007	Adm95007 Inverted
c	12	19	100.0	19	12	ADQ16519	Adq16519 Transposo
	13	19	100.0	19	12	ADQ16518	Adq16518 Transposo
	14	19	100.0	31	10	ADF79375	Adf79375 Transposo
c	15	19	100.0	32	6	ABK87204	Abk87204 Synthetic
	16	19	100.0	32	10	ADF79376	Adf79376 Transposo
c	17	19	100.0	38	6	ABK87203	Abk87203 Synthetic
	18	19	100.0	39	10	AAD58905	Aad58905 Tn5 Trans
c	19	19	100.0	39	10	AAD58905	Aad58905 Tn5 Trans
	20	19	100.0	41	10	ADC53953	Adc53953 Rhodococc
	21	19	100.0	41	10	ADG20413	Adg20413 Pseudomon
	22	19	100.0	41	12	ADF72744	Adf72744 pMOD1 tra
	23	19	100.0	41	12	ADM35526	Adm35526 Kap1 prom
	24	19	100.0	41	13	ADR44424	Adr44424 Plasmid p
	25	19	100.0	42	10	ADC53954	Adc53954 Rhodococc
	26	19	100.0	42	10	ADG20414	Adg20414 Pseudomon
	27	19	100.0	42	12	ADF72745	Adf72745 pMOD1 tra
	28	19	100.0	42	12	ADM35527	Adm35527 Kap1 prom
	29	19	100.0	42	13	ADR44425	Adr44425 Plasmid p
	30	19	100.0	80	10	ACH00838	Ach00838 Primer 98
	31	19	100.0	84	12	ADH48035	Adh48035 Transposo
	32	19	100.0	85	10	ACH00837	Ach00837 Primer 97
c	33	19	100.0	94	12	ADH48036	Adh48036 Transposo
	34	19	100.0	136	10	AAD58813	Aad58813 Transposo
c	35	19	100.0	136	10	AAD58813	Aad58813 Transposo
	36	19	100.0	137	10	AAD58812	Aad58812 Transposo
c	37	19	100.0	137	10	AAD58812	Aad58812 Transposo
	38	19	100.0	137	10	AAD58811	Aad58811 Transposo
c	39	19	100.0	137	10	AAD58811	Aad58811 Transposo
c	40	19	100.0	160	5	ABA06312	Aba06312 Soy bean
	41	19	100.0	171	4	AAC91698	Aac91698 Ovalbumin
c	42	19	100.0	171	4	AAC91698	Aac91698 Ovalbumin
	43	19	100.0	204	4	AAC91699	Aac91699 I-Ab epit
c	44	19	100.0	204	4	AAC91699	Aac91699 I-Ab epit
	45	19	100.0	758	6	AAD35078	Aad35078 MOD3-pTac

ALIGNMENTS

RESULT 1
AAV28400
ID AAV28400 standard; DNA; 19 BP.
XX
AC AAV28400;
XX
DT 24-JUL-1998 (first entry)
XX
DE Transposon 5 (Tn5) mutant outside end (OE) sequence 1.
XX
KW Tn5 transposase; modified; enzyme; in vitro transposition; mutant;
KW target; marker; transposon 5; plasmid pRZTL1; outside end; OE; ds.
XX
OS Escherichia coli.
XX
PN WO9810077-A1.
XX
PD 12-MAR-1998.
XX
PF 09-SEP-1997; 97WO-US015941.
XX
PR 09-SEP-1996; 96US-00814877.
PR 02-MAY-1997; 97US-00850880.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Reznikoff WS, Goryshin IY, Zhou H;
XX
DR WPI; 1998-193627/17.
XX
PT Modified Tn5 transposase construct used in novel system for in vitro
PT transposition - used to, e.g. create absolute defective mutants, provide
PT selective markers and to facilitate insertion of specialised DNA
PT sequences into target DNA.
XX
PS Claim 13; Page 55; 73pp; English.
XX
CC This is the transposon 5 (Tn5) mutant outside end (OE) sequence used in
CC the novel genetic construct of the invention. The genetic construct
CC comprises a nucleotide sequence encoding a modified Tn5 transposase
CC enzyme that has both greater avidity for Tn5 OE repeats and is less
CC likely to assume an inactive multimeric form than a wild type Tn5
CC transposase and a transposable DNA sequence flanked at its 5' and 3' ends

CC by an 18 or 19 base pair flanking DNA sequence comprising nucleotide A at
CC position 10, T at 11 and A at 12. The modified Tn5 transposase and the
CC transposable DNA which is a DNA donor molecule are used in a system for
CC in vitro transposition. The system and method can be used to create
CC absolute defective mutants, to provide selective markers to target DNA,
CC to provide portable regions of homology to a target DNA, to facilitate
CC insertion of specialised DNA sequences into target DNA, to provide primer
CC binding sites or tags for DNA sequencing , to facilitate production of
CC genetic fusion for gene expression studies and protein domain mapping, as
CC well as to bring together other desired combinations of DNA sequences
CC (combinatorial genetics). The modified Tn5 transposase facilitates in
CC vitro transposition reaction rates of at least about 100-fold higher than
CC can be achieved using wild type transposase (as measure in vivo). In
CC vitro transposition using this system can also use donor DNA and target
CC DNA that is circular or linear. The system also requires no outside high
CC energy source and no other protein other than the modified transposase
XX
SQ Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 1 CTGTCTCTTATACACATCT 19

RESULT 2
AAZ06436
ID AAZ06436 standard; DNA; 19 BP.
XX
AC AAZ06436;
XX
DT 09-NOV-1999 (first entry)
XX
DE Mutant Outside End (OE) termini 4/17/18.
XX
KW transposase; modified form; wildtype; multimeric; OE termini; IE termini;
KW outside end termini; inside end termini; plasmid; repeat sequence;
KW mutation; ds.
XX
OS Transposon Tn5.
OS Synthetic.
XX
PN US5948622-A.
XX
PD 07-SEP-1999.
XX
PF 06-OCT-1997; 97US-00944916.
XX
PR 09-SEP-1996; 96US-00814877.
PR 02-MAY-1997; 97US-00850880.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Zhou H, York DL, Goryshin IY, Reznikoff WS;
XX
DR WPI; 1999-517947/43.
XX
PT In vitro transposition using a Tn5 based genetic construct.
XX
PS Example 1; Col 45; 48pp; English.
XX
CC This is the nucleotide sequence of a mutant Outside End termini, which
CC differs from the wildtype sequence at positions 4, 17 and 18, counting
CC from the 5' end. Wildtype Outside End (OE, AAZ06435) and Inside End (IE,
CC AAZ06438) were compared and an effort made to randomize the nucleotides
CC at each of the seven positions of difference. A population of
CC oligonucleotides degenerate at each position of difference was created.
CC This resulted in individual oligonucleotides in the population randomly
CC included either the nucleotide of the wildtype OE or the wildtype IE. 128
CC distinct oligonucleotides were generated, which had the sequence
CC characteristics of both OE and IE and so can be referred to as OE/IE-like
CC sequences. Two of these OE/IE-like sequences are the mutant OE sequences
CC AAZ06436 and AAZ06437
XX
SQ Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 1 CTGTCTCTTATACACATCT 19

RESULT 3
AAA11737
ID AAA11737 standard; DNA; 19 BP.
XX
AC AAA11737;
XX
DT 21-JUL-2000 (first entry)
XX
DE Transposon Tn5 interacting nucleotide #1.
XX
KW Transposon; transposase; insertion mutation; synaptic complex; ss.
XX
OS Transposon Tn5.
XX
PN WO200017343-A1.
XX
PD 30-MAR-2000.
XX
PF 21-SEP-1999; 99WO-US021960.
XX
PR 23-SEP-1998; 98US-00159363.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Reznikoff WS, Goryshin IY;
XX
DR WPI; 2000-283573/24.
XX
PT Making insertional mutations at random or quasi-random positions in
PT cellular nucleic acids in target cells, useful for identifying
PT chromosomal regions involved in expressing or regulating expression of
PT proteins.
XX
PS Claim 13; Page 17; 25pp; English.
XX
CC This invention describes a novel method (I) for making an insertional
CC mutation at a random or quasi-random position in cellular nucleic acid in
CC a target cell. The invention describes a method (II) for forming a
CC synaptic complex between a Tn5 transposase protein (X) and a
CC polynucleotide (Y) that comprises a pair of nucleotide sequences adapted
CC for operably interacting with the Tn5 transposase to form a synaptic
CC complex and a transposable nucleotide sequence between them, comprising
CC combining (X) and (Y) in vitro under conditions that disfavor
CC polynucleotide strand transfer to form the synaptic complex. Methods for
CC the insertion of exogenous nucleic acids into the nucleic acids of target
CC cells are used to identify chromosomal regions involved in expressing or
CC regulating expression of proteins. The same methods may be used in the
CC development of new therapeutic agents. The transposable polynucleotides
CC used to form synaptic complexes can consist of transposon apart from any
CC flanking sequences. This is advantageous in that it reduces the
CC likelihood of intramolecular transposition and increases the likelihood
CC of transposition into a target genome. Eliminating donor backbone
CC sequences from the polynucleotide simplifies preparation of the
CC transposon sequences to be used in (I). Additionally, the synaptic
CC complex can form under conditions that disfavor non-productive
CC intramolecular transposition events. This is advantageous because all of
CC the synaptic complexes can undergo transposition when combined with
CC cellular DNA. Little, if any, of the nucleic acid in the synaptic

```
CC   complexes is inactive. This sequence represents a novel transposon Tn5
CC   transposase interacting nucleotide fragment described in the method of
CC   the invention
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 4
AAD21282
ID   AAD21282 standard; DNA; 19 BP.
XX
AC   AAD21282;
XX
DT   28-JAN-2002  (first entry)
XX
DE   Mosaic terminal sequence #1.
XX
KW   Insertional mutation; synaptic complex; transposon; screening;
KW   mosaic terminus; ds.
XX
OS   Unidentified.
XX
PN   US6294385-B1.
XX
PD   25-SEP-2001.
XX
PF   10-AUG-2000; 2000US-00635969.
XX
PR   23-SEP-1998;   98US-00159363.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Goryshin IY,  Reznikoff WS;
XX
DR   WPI; 2001-656171/75.
XX
PT   Making an insertional mutations, especially useful for efficiently
PT   inserting a transposable polynucleotide in a target cell, comprises
PT   introducing into the target cell a synaptic complex.
XX
PS   Disclosure; Fig 3; 11pp; English.
XX
CC   The present invention relates to a method for making an insertional
CC   mutation at a random or quasi-random position in cellular nucleic acid in
CC   a target cell comprising introducing into the target cell a synaptic
CC   complex. The method is particularly useful for efficiently inserting a
CC   transposable polynucleotide at random or quasi-random locations in the
CC   chromosomal or extra-chromosomal nucleic acid of a target cell. The
CC   method may also be used for screening the genome of cells that comprise
CC   an insertional mutation that induces a phenotypic or genotypic change
CC   relative to the cells that are not subject to insertional mutagenesis.
CC   The present sequence is the mosaic terminus sequence which is used to
CC   enhance the transposition frequency over that of wild-type Tn5
CC   transposon, used in the exemplification of the invention
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 5
AAC91688
ID   AAC91688 standard; DNA; 19 BP.
XX
AC   AAC91688;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Transposon Tn5 mosaic insertion end.
XX
KW   Transposable element; MHC epitope; major histocompatibility complex;
KW   intracellular bacterial pathogen; loxP site; Cre recombinase;
KW   insertion end; in-frame fusion; detection; antigen;
KW   disseminated insertions of class-I epitopes; DICE-I; transposon Tn5;
KW   mosaic insertion end; ds.
XX
OS   Escherichia coli.
OS   Synthetic.
XX
PN   WO200071158-A1.
XX
PD   30-NOV-2000.
XX
PF   26-MAY-2000; 2000WO-US014687.
XX
PR   26-MAY-1999;   99US-0136210P.
XX
PA   (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI   Heffron FL,  Parker DC,  Ellefson DD;
XX
DR   WPI; 2001-031967/04.
XX
PT   Transposable element for detecting an antigenic epitope of a pathogen,
PT   comprising 5' and 3' recombining sites, nucleic acid sequences encoding a
PT   selectable marker and major histocompatibility complex (MHC) epitope, and
PT   an insertion end.
XX
PS   Claim 17; Fig 11; 63pp; English.
XX
CC   The invention relates to a novel transposable element comprising DNA
CC   encoding a selectable marker (e.g., antibiotic resistance) located
CC   between a 5' recombining site and a 3' recombining site (e.g., loxP
CC   sites); DNA encoding an MHC (major histocompatibility complex) epitope
CC   either 5' of the 5' recombining site or 3' of the 3' recombining site;
CC   and insertion ends comprising an inverted repeat sequence at the 5' and
CC   3' ends of the transposable element sufficient for integration of the
CC   transposable element. The transposable elements of the invention are able
CC   to introduce in-frame insertions throughout the chromosome of an
CC   intracellular bacterial pathogen. This system "tags" the bacterial gene
CC   and resulting protein, allowing the identification of proteins secreted
CC   across the membranes of the eukaryotic cell infected by the bacterium. In
CC   one embodiment, the transposable elements contain an antibiotic
CC   resistance cassette, two minimal loxP recombination sites, an MHC class I
CC   or class II epitope, and flanking insertion ends. A transposase, such as
CC   the Cre recombinase protein, is expressed in trans from a plasmid, or can
CC   be included in the transposable element. The Cre recombinase loops out
CC   the intervening sequences containing the antibiotic resistance cassette.
CC   When the transposable element inserts within a gene, the resolved
CC   insertion places the MHC class I or class II epitope in frame with the
CC   gene. The transposable elements of the invention are useful for detecting
CC   an antigenic epitope of an intracellular bacterial pathogen, such as
CC   Salmonella sp., Mycobacterium tuberculosis and Listeria monocytogenes.
CC   Certain embodiments of the technology, termed "disseminated insertions of
CC   class-I epitopes" (DICE-I; DICE-II for class II epitopes) allow the rapid
CC   and accurate identification of proteins involved in bacterial
CC   pathogenesis so that such proteins can be used as vaccine and drug
CC   targets. Carrier vaccines may be generated by infecting bacteria with a
CC   transposable element of the invention which additionally comprises an
CC   antigen associated with a disease, preferably cancer or a viral or
CC   bacterial disease, operably linked to the MHC epitope DNA of the
CC   transposable element. The present sequence represents a transposon Tn5
```

```
CC   mosaic insertion end claimed for use in a transposable element of the
CC   invention
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 4;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 6
ADA13512
ID   ADA13512 standard; DNA; 19 BP.
XX
AC   ADA13512;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Outer end transposase recognition sequence SEQ ID NO:3.
XX
KW   transposon; TnKGloxP; outer end transposase recognition sequence;
KW   OE sequence; loxP site; KmR; kanamycin resistance gene; GFP;
KW   green fluorescent protein; CmR; chloramphenicol resistance gene; gene;
KW   ds.
XX
OS   Synthetic.
XX
PN   WO2003070955-A1.
XX
PD   28-AUG-2003.
XX
PF   31-OCT-2002; 2002WO-KR002033.
XX
PR   22-FEB-2002; 2002KR-00009647.
XX
PA   (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Kim S,  Yu B;
XX
DR   WPI; 2003-679884/64.
XX
PT   New transposon TnKGloxP comprising outer end transposase recognition
PT   sequences with base sequence on one end, reverse-complementary sequence
PT   on the other end, loxP site, KmR and GFP gene useful for deleting
PT   chromosome specific sites.
XX
PS   Claim 1; Page 3; 38pp; English.
XX
CC   The present invention describes a transposon TnKGloxP comprising outer
CC   end transposase recognition sequences (OE sequence) having a 19 base pair
CC   sequence (see ADA13512) on one end, its reverse-complementary sequence on
CC   the other end, loxP site expressed as a 34 base pair sequence (see
CC   ADA13513), KmR (kanamycin resistance) gene expressed as a 996 base pair
CC   sequence (see ADA13514) and GFP (green fluorescent protein) gene
CC   expressed as a 947 base pair sequence (see ADA13515). Also described: (1)
CC   a transposon TnCloxP comprising OE sequences having (see ADA13512), its
CC   reverse-complementary sequence of the other end, the loxP site (see
CC   ADA13513) and the CmR (chloramphenicol resistance) gene expressed as a
CC   1069 base pair sequence (see ADA13516); and (2) constructing novel
CC   strains containing deletion of a specific chromosomal site, comprising:
CC   (a) preparing two transposons comprising outer end transposase
CC   recognition sequences, loxP site and different selectable markers; (b)
CC   inserting the two transposons, respectively, into random positions of
CC   different microbial chromosomes and determining the each inserted sites;
CC   (c) integrating the two microbial chromosomes by P1 phage transduction to
CC   position the two transposons comprising different selectable markers on
CC   one chromosome; and (d) deleting a chromosomal site between the two loxP
CC   sites by expressing Cre gene through Cre expression vector introduced.
CC   The transposon is useful for deleting chromosome specific sites. The
CC   vector is useful for the preparation of the transposon. The present
CC   sequence represents an OE sequence from the present invention.
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 9;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 7
AAD58809
ID   AAD58809 standard; DNA; 19 BP.
XX
AC   AAD58809;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Tn5 transposon mosaic element (ME) DNA #1.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME; ds.
XX
OS   Unidentified.
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Disclosure; Page 4; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is Tn5
CC   transposon mosaic element (ME) DNA. This sequence is used to illustrate
CC   the method of the invention
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 10;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19
```

```
RESULT 8
AAD58810/c
ID   AAD58810 standard; DNA; 19 BP.
XX
AC   AAD58810;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Tn5 transposon mosaic element (ME) DNA #2.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME; ds.
XX
OS   Unidentified.
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 12; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is Tn5
CC   transposon mosaic element (ME) DNA. This sequence is used to illustrate
CC   the method of the invention
XX
SQ   Sequence 19 BP; 8 A; 1 C; 6 G; 4 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         19 CTGTCTCTTATACACATCT 1


RESULT 9
ADF79368
ID   ADF79368 standard; DNA; 19 BP.
XX
AC   ADF79368;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Transposon Tn5 transposase recognition site.
XX
KW   Transposon Tn5; Transposon TnRIBD; genome deletion; non-essential gene;
KW   transposase; ds.
XX
OS   Chimeric.
OS   Escherichia coli.
XX
PN   WO2003089639-A1.
XX
PD   30-OCT-2003.
XX
PF   18-APR-2003; 2003WO-KR000798.
XX
PR   20-APR-2002; 2002KR-00021811.
XX
PA   (KOAD ) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI   Kim S,  Sung B,  Yu B,  Kim J,  Lee W,  Lee C,  Lee J;
XX
DR   WPI; 2003-854122/79.
XX
PT   New transposon, useful for developing a mutant strain with deletion of an
PT   optional part of the chromosome, and for identifying non-essential genes
PT   for growth of microorganisms.
XX
PS   Claim 1; SEQ ID NO 2; 47pp; English.
XX
CC   The present sequence is that of transposon Tn5 transposase recognition
CC   site (outer element). This recognition site was used in the construction
CC   of transposon TnRIBD ADF79367. A claimed method for developing a mutant
CC   microbial strain with deletion of an optional part of the chromosome
CC   involves inserting transposon TnRIBD into an optional site, identifying
CC   the insertion site, and deleting the parts of the chromosome on the left
CC   and right hand sides of the insertion site using a transposase expression
CC   vector. A claimed method for identifying non-essential genes for growth
CC   involves constructing a new mutant strain by deleting an optional part of
CC   the genome, identifying the genes in the deleted part, and investigating
CC   the survival of the mutant strain. The methods can be used to develop
CC   novel strains of Escherichia coli and other microorganisms.
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 10
ADM95006
ID   ADM95006 standard; DNA; 19 BP.
XX
AC   ADM95006;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Inverted repeat sequence, SEQ ID 1.
XX
KW   Inverted repeat sequence; transposable element; transposon Tn5; ds.
XX
OS   Synthetic.
XX
PN   CA2396611-A1.
XX
PD   31-JAN-2004.
XX
PF   31-JUL-2002; 2002CA-02396611.
XX
PR   31-JUL-2002; 2002CA-02396611.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dyson PJ,  Herron P;
```

```
XX
DR   WPI; 2004-192322/19.
XX
PT   New nucleic acid construct comprising inverted repeat sequences of a
PT   transposable element and an origin of transfer between the inverted
PT   repeat sequences, useful for introducing genetic disruptions in a
PT   bacterial genetic material.
XX
PS   Claim 4; SEQ ID NO 1; 46pp; English.
XX
CC   The present invention relates to a nucleic acid construct (I), which
CC   comprises inverted repeat sequences (ADM95006-ADM95009 and ADM95017-
CC   ADM95018) of a transposable element and an origin of transfer that lies
CC   between the inverted repeat sequences, such that a transposition event
CC   involving the inverted repeat sequences will result in the origin of
CC   transfer being included in the resultant insertion at the transposition
CC   target site. Preferably the inverted repeat sequences are or are derived
CC   from the OE and/or IE inverted repeat sequences of the transposon Tn5.
CC   The origin of transfer is an oriT, which can be mobilized by the helper
CC   plasmids pUZ8002 and pUB307, and has a sequence of ADM95010. The
CC   construct comprises a promoterless reporter gene located between the
CC   inverted repeat sequences, where the promoterless reporter gene is
CC   operatively associated with a ribosome binding site, and the construct
CC   further comprises upstream of the reporter gene and ribosome binding site
CC   and between the inverted repeat sequences, a translational stop sequence.
CC   The construct lacks an origin of replication, is linear, and consists
CC   essentially of the inverted repeat sequences and any sequences located
CC   between. The nucleic acid construct is useful for introducing genetic
CC   disruptions in a bacterial genetic material, particularly that of the
CC   Streptomyces species.
XX
SQ   Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 12;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 11
ADM95007/c
ID   ADM95007 standard; DNA; 19 BP.
XX
AC   ADM95007;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Inverted repeat sequence, SEQ ID 2.
XX
KW   Inverted repeat sequence; transposable element; transposon Tn5; ds.
XX
OS   Synthetic.
XX
PN   CA2396611-A1.
XX
PD   31-JAN-2004.
XX
PF   31-JUL-2002; 2002CA-02396611.
XX
PR   31-JUL-2002; 2002CA-02396611.
XX
PA   (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI   Dyson PJ,  Herron P;
XX
DR   WPI; 2004-192322/19.
XX
PT   New nucleic acid construct comprising inverted repeat sequences of a
PT   transposable element and an origin of transfer between the inverted
PT   repeat sequences, useful for introducing genetic disruptions in a
PT   bacterial genetic material.
XX
PS   Claim 4; SEQ ID NO 2; 46pp; English.
XX
CC   The present invention relates to a nucleic acid construct (I), which
CC   comprises inverted repeat sequences (ADM95006-ADM95009 and ADM95017-
CC   ADM95018) of a transposable element and an origin of transfer that lies
CC   between the inverted repeat sequences, such that a transposition event
CC   involving the inverted repeat sequences will result in the origin of
CC   transfer being included in the resultant insertion at the transposition
CC   target site. Preferably the inverted repeat sequences are or are derived
CC   from the OE and/or IE inverted repeat sequences of the transposon Tn5.
CC   The origin of transfer is an oriT, which can be mobilized by the helper
CC   plasmids pUZ8002 and pUB307, and has a sequence of ADM95010. The
CC   construct comprises a promoterless reporter gene located between the
CC   inverted repeat sequences, where the promoterless reporter gene is
CC   operatively associated with a ribosome binding site, and the construct
CC   further comprises upstream of the reporter gene and ribosome binding site
CC   and between the inverted repeat sequences, a translational stop sequence.
CC   The construct lacks an origin of replication, is linear, and consists
CC   essentially of the inverted repeat sequences and any sequences located
CC   between. The nucleic acid construct is useful for introducing genetic
CC   disruptions in a bacterial genetic material, particularly that of the
CC   Streptomyces species.
XX
SQ   Sequence 19 BP; 8 A; 1 C; 6 G; 4 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 12;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         19 CTGTCTCTTATACACATCT 1


RESULT 12
ADQ16519/c
ID   ADQ16519 standard; DNA; 19 BP.
XX
AC   ADQ16519;
XX
DT   23-SEP-2004  (first entry)
XX
DE   Transposon Tn5 mosaic element #2.
XX
KW   Transposon Tn5; ss; transposase mediated integration; transposon;
KW   transposase; Tn5 mosaic element; random insertional mutagenesis;
KW   RNA polymerase III promoter; U1 snRNA gene.
XX
OS   Transposon Tn5.
XX
PN   US2004126887-A1.
XX
PD   01-JUL-2004.
XX
PF   08-NOV-2002; 2002US-00291342.
XX
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2004-542387/52.
XX
PT   Composition useful for enhancing transposase mediated integration of
PT   transposon into target nucleic acid, comprising integrator complex, and
PT   enhancing reagent.
XX
PS   Example; SEQ ID NO 4; 14pp; English.
```

XX
CC The invention relates to a composition for enhancing transposase mediated
CC integration of a transposon into a target nucleic acid, comprising an
CC integrator complex and an enhancing reagent. The invention also relates
CC to a method of integrating a nucleic acid into a target nucleic acid,
CC involving making a transposon, forming an integrator complex, combining
CC the integrator complex and a cationic enhancing reagent together in
CC solution, and incubating the composition with a target nucleic acid,
CC where the transposase integrates the transposon into the target nucleic
CC acid. The transposase is a hyperactive mutant Tn5 transposase. The Tn5
CC transposase is flanked by elements chosen from Tn5 outer elements, Tn5
CC inner elements and Tn5 mosaic elements. The enhancing reagent is chosen
CC from transfection reagents, polycations, cationic polymers and cationic
CC lipids. The enhancing reagent comprises both cationic proteins and
CC cationic lipids. The composition and the method are useful for providing
CC random insertional mutagenesis, in which integration of a transposon into
CC a target nucleic acid inserts a molecular tag or disrupts a target
CC sequence, where the integration of a molecular tag facilitates cloning,
CC sequencing or identification by providing a detectable marker, and the
CC integration into a coding region disrupts gene function and facilitates
CC study of a gene. The composition is useful for identifying enhancer
CC elements, for sequencing DNA and for integrating large DNA fragments with
CC known ends into a target nucleic acid such as a plasmid, an artificial
CC chromosome or a viral vector. The composition is also useful for
CC integrating e.g. therapeutic genes, siRNA genes, reporter genes, marker
CC or tag sequences, genes containing RNA polymerase III promoters or
CC modified U1 snRNA genes. This sequence represents a transposon Tn5 mosaic
CC element used in the scope of the invention.
XX
SQ Sequence 19 BP; 8 A; 1 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 19 CTGTCTCTTATACACATCT 1

RESULT 13
ADQ16518
ID ADQ16518 standard; DNA; 19 BP.
XX
AC ADQ16518;
XX
DT 23-SEP-2004 (first entry)
XX
DE Transposon Tn5 mosaic element #1.
XX
KW Transposon Tn5; ss; transposase mediated integration; transposon;
KW transposase; Tn5 mosaic element; random insertional mutagenesis;
KW RNA polymerase III promoter; U1 snRNA gene.
XX
OS Transposon Tn5.
XX
PN US2004126887-A1.
XX
PD 01-JUL-2004.
XX
PF 08-NOV-2002; 2002US-00291342.
XX
PR 08-NOV-2001; 2001US-0344865P.
XX
PA (WOOD/) WOODDELL C.
PA (HERW/) HERWEIJER H.
PA (WOLF/) WOLFF J A.
XX
PI Wooddell C, Herweijer H, Wolff JA;
XX
DR WPI; 2004-542387/52.
XX
PT Composition useful for enhancing transposase mediated integration of
PT transposon into target nucleic acid, comprising integrator complex, and
PT enhancing reagent.
XX
PS Example; SEQ ID NO 3; 14pp; English.
XX
CC The invention relates to a composition for enhancing transposase mediated
CC integration of a transposon into a target nucleic acid, comprising an
CC integrator complex and an enhancing reagent. The invention also relates
CC to a method of integrating a nucleic acid into a target nucleic acid,
CC involving making a transposon, forming an integrator complex, combining
CC the integrator complex and a cationic enhancing reagent together in
CC solution, and incubating the composition with a target nucleic acid,
CC where the transposase integrates the transposon into the target nucleic
CC acid. The transposase is a hyperactive mutant Tn5 transposase. The Tn5
CC transposase is flanked by elements chosen from Tn5 outer elements, Tn5
CC inner elements and Tn5 mosaic elements. The enhancing reagent is chosen
CC from transfection reagents, polycations, cationic polymers and cationic
CC lipids. The enhancing reagent comprises both cationic proteins and
CC cationic lipids. The composition and the method are useful for providing
CC random insertional mutagenesis, in which integration of a transposon into
CC a target nucleic acid inserts a molecular tag or disrupts a target
CC sequence, where the integration of a molecular tag facilitates cloning,
CC sequencing or identification by providing a detectable marker, and the
CC integration into a coding region disrupts gene function and facilitates
CC study of a gene. The composition is useful for identifying enhancer
CC elements, for sequencing DNA and for integrating large DNA fragments with
CC known ends into a target nucleic acid such as a plasmid, an artificial
CC chromosome or a viral vector. The composition is also useful for
CC integrating e.g. therapeutic genes, siRNA genes, reporter genes, marker
CC or tag sequences, genes containing RNA polymerase III promoters or
CC modified U1 snRNA genes. This sequence represents a transposon Tn5 mosaic
CC element used in the scope of the invention.
XX
SQ Sequence 19 BP; 4 A; 6 C; 1 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 1 CTGTCTCTTATACACATCT 19

RESULT 14
ADF79375
ID ADF79375 standard; DNA; 31 BP.
XX
AC ADF79375;
XX
DT 26-FEB-2004 (first entry)
XX
DE Transposon Tn5 transposase recognition site 5' PCR primer.
XX
KW Transposon TnRIBD; Transposon Tn5; genome deletion; non-essential gene;
KW transposase; PCR; primer; ss.
XX
OS Chimeric.
OS Escherichia coli.
XX
PN WO2003089639-A1.
XX
PD 30-OCT-2003.
XX
PF 18-APR-2003; 2003WO-KR000798.
XX
PR 20-APR-2002; 2002KR-00021811.
XX
PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI Kim S, Sung B, Yu B, Kim J, Lee W, Lee C, Lee J;
XX
DR WPI; 2003-854122/79.

XX
PT New transposon, useful for developing a mutant strain with deletion of an
PT optional part of the chromosome, and for identifying non-essential genes
PT for growth of microorganisms.
XX
PS Example 1; SEQ ID NO 9; 47pp; English.
XX
CC The present sequence is that of a 5' PCR primer for the transposon Tn5
CC transposase recognition site ADF79368. It was used with a 3' primer
CC ADF79376 to amplify the recognition site from pMb2. The recognition site
CC was used in the construction of transposon TnRIBD ADF79367. A claimed
CC method for developing a mutant microbial strain with deletion of an
CC optional part of the chromosome involves inserting transposon TnRIBD into
CC an optional site, identifying the insertion site, and deleting the parts
CC of the chromosome on the left and right hand sides of the insertion site
CC using a transposase expression vector. A claimed method for identifying
CC non-essential genes for growth involves constructing a new mutant strain
CC by deleting an optional part of the genome, identifying the genes in the
CC deleted part, and investigating the survival of the mutant strain. The
CC methods can be used to develop novel strains of Escherichia coli and
CC other microorganisms.
XX
SQ Sequence 31 BP; 7 A; 9 C; 4 G; 11 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 12 CTGTCTCTTATACACATCT 30

RESULT 15
ABK87204/c
ID ABK87204 standard; DNA; 32 BP.
XX
AC ABK87204;
XX
DT 24-SEP-2002 (first entry)
XX
DE Synthetic compressed transposase-binding linker B.
XX
KW Transposase-interacting inverted repeat sequence pair;
KW transposase enzyme; gene fusion library; transposable element;
KW transposase-binding linker; ds.
XX
OS Synthetic.
XX
PN WO200246444-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US046311.
XX
PR 05-DEC-2000; 2000US-0251482P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Goryshin IY, Naumann TA, Reznikoff WS;
XX
DR WPI; 2002-527923/56.
XX
PT Transposable polynucleotide for manipulating nucleic acids to produce
PT gene fusions, comprises two or more transposase-interacting inverted
PT repeat sequence pairs.
XX
PS Disclosure; Fig 1; 53pp; English.
XX
CC The present invention relates to a new polynucleotide comprising distinct
CC first and second transposase-interacting inverted repeat sequence pairs.
CC Each pair has a specificity for binding to and interacting with a
CC distinct transposase enzyme, members of the first sequence pair flanking
CC members of the second sequence pair. The invention is useful for
CC producing a gene fusion library and is also useful for deleting a portion
CC of a chromosome and for cloning a portion of a chromosome of a host cell.
CC The invention is further useful for inserting a preselected
CC polynucleotide sequence insert into a chromosome of a host cell .
CC Transposition occurs without regard to the sequences of the nucleic acid
CC into which the transposable elements transpose. Large libraries having a
CC high level of variability can be produced using the polynucleotide of the
CC invention. The present nucleic acid sequence represents the compressed
CC transposase-binding linker B sequence that is part of a transposase-
CC interacting inverted repeat sequence pair, as described above
XX
SQ Sequence 32 BP; 9 A; 5 C; 8 G; 10 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 32 CTGTCTCTTATACACATCT 14

RESULT 16
ADF79376
ID ADF79376 standard; DNA; 32 BP.
XX
AC ADF79376;
XX
DT 26-FEB-2004 (first entry)
XX
DE Transposon Tn5 transposase recognition site 3' PCR primer.
XX
KW Transposon TnRIBD; Transposon Tn5; genome deletion; non-essential gene;
KW transposase; PCR; primer; ss.
XX
OS Chimeric.
OS Escherichia coli.
XX
PN WO2003089639-A1.
XX
PD 30-OCT-2003.
XX
PF 18-APR-2003; 2003WO-KR000798.
XX
PR 20-APR-2002; 2002KR-00021811.
XX
PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
XX
PI Kim S, Sung B, Yu B, Kim J, Lee W, Lee C, Lee J;
XX
DR WPI; 2003-854122/79.
XX
PT New transposon, useful for developing a mutant strain with deletion of an
PT optional part of the chromosome, and for identifying non-essential genes
PT for growth of microorganisms.
XX
PS Example 1; SEQ ID NO 10; 47pp; English.
XX
CC The present sequence is that of a 3' PCR primer for the transposon Tn5
CC transposase recognition site ADF79368. It was used with a 5' primer
CC ADF79375 to amplify the recognition site from pMb2. The recognition site
CC was used in the construction of transposon TnRIBD ADF79367. A claimed
CC method for developing a mutant microbial strain with deletion of an
CC optional part of the chromosome involves inserting transposon TnRIBD into
CC an optional site, identifying the insertion site, and deleting the parts
CC of the chromosome on the left and right hand sides of the insertion site
CC using a transposase expression vector. A claimed method for identifying
CC non-essential genes for growth involves constructing a new mutant strain
CC by deleting an optional part of the genome, identifying the genes in the
CC deleted part, and investigating the survival of the mutant strain. The
CC methods can be used to develop novel strains of Escherichia coli and
CC other microorganisms.

```
XX
SQ   Sequence 32 BP; 7 A; 10 C; 4 G; 11 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 19;  DB 10;  Length 32;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         13 CTGTCTCTTATACACATCT 31


RESULT 17
ABK87203/c
ID   ABK87203 standard; DNA; 38 BP.
XX
AC   ABK87203;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Synthetic full-length transposase-binding linker B.
XX
KW   Transposase-interacting inverted repeat sequence pair;
KW   transposase enzyme; gene fusion library; transposable element;
KW   transposase-binding linker; ds.
XX
OS   Synthetic.
XX
PN   WO200246444-A2.
XX
PD   13-JUN-2002.
XX
PF   05-DEC-2001; 2001WO-US046311.
XX
PR   05-DEC-2000; 2000US-0251482P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Goryshin IY,  Naumann TA,  Reznikoff WS;
XX
DR   WPI; 2002-527923/56.
XX
PT   Transposable polynucleotide for manipulating nucleic acids to produce
PT   gene fusions, comprises two or more transposase-interacting inverted
PT   repeat sequence pairs.
XX
PS   Disclosure; Fig 1; 53pp; English.
XX
CC   The present invention relates to a new polynucleotide comprising distinct
CC   first and second transposase-interacting inverted repeat sequence pairs.
CC   Each pair has a specificity for binding to and interacting with a
CC   distinct transposase enzyme, members of the first sequence pair flanking
CC   members of the second sequence pair. The invention is useful for
CC   producing a gene fusion library and is also useful for deleting a portion
CC   of a chromosome and for cloning a portion of a chromosome of a host cell.
CC   The invention is further useful for inserting a preselected
CC   polynucleotide sequence insert into a chromosome of a host cell .
CC   Transposition occurs without regard to the sequences of the nucleic acid
CC   into which the transposable elements transpose. Large libraries having a
CC   high level of variability can be produced using the polynucleotide of the
CC   invention. The present nucleic acid sequence represents the full-length
CC   transposase-binding linker B sequence that is part of a transposase-
CC   interacting inverted repeat sequence pair, as described above
XX
SQ   Sequence 38 BP; 11 A; 6 C; 9 G; 12 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 19;  DB 6;  Length 38;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         38 CTGTCTCTTATACACATCT 20


RESULT 18
AAD58905
ID   AAD58905 standard; DNA; 39 BP.
XX
AC   AAD58905;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Tn5 Transposon containing mosaic element (ME).
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   1. .39
FT                   /*tag=  a
FT                   /rpt_type= INVERTED
FT   repeat_unit     1. .19
FT                   /*tag=  b
FT   misc_feature    20
FT                   /*tag=  c
FT                   /note= "This base is shown as (N)x which represents a
FT                   sequence that is inserted between the flanking mosaic
FT                   sequences"
FT   repeat_unit     21. .39
FT                   /*tag=  d
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 4; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is Tn5
CC   transposon containing mosaic element (ME). This sequence is used to
CC   illustrate the method of the invention
XX
SQ   Sequence 39 BP; 12 A; 7 C; 7 G; 12 T; 0 U; 1 Other;

  Query Match           100.0%;  Score 19;  DB 10;  Length 39;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
```

```
Db          1 CTGTCTCTTATACACATCT 19


RESULT 19
AAD58905/c
ID   AAD58905 standard; DNA; 39 BP.
XX
AC   AAD58905;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Tn5 Transposon containing mosaic element (ME).
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   1. .39
FT                   /*tag=  a
FT                   /rpt_type= INVERTED
FT   repeat_unit     1. .19
FT                   /*tag=  b
FT   misc_feature    20
FT                   /*tag=  c
FT                   /note= "This base is shown as (N)x which represents a
FT                   sequence that is inserted between the flanking mosaic
FT                   sequences"
FT   repeat_unit     21. .39
FT                   /*tag=  d
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 4; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is Tn5
CC   transposon containing mosaic element (ME). This sequence is used to
CC   illustrate the method of the invention
XX
SQ   Sequence 39 BP; 12 A; 7 C; 7 G; 12 T; 0 U; 1 Other;

  Query Match           100.0%;  Score 19;  DB 10;  Length 39;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         39 CTGTCTCTTATACACATCT 21


RESULT 20
ADC53953
ID   ADC53953 standard; DNA; 41 BP.
XX
AC   ADC53953;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Rhodococcus promoter related primer #SEQ ID 2.
XX
KW   Promoter; protein production; bacterial; PCR; primer; ss.
XX
OS   Rhodococcus erythropolis.
XX
PN   JP2003144166-A.
XX
PD   20-MAY-2003.
XX
PF   14-NOV-2001; 2001JP-00348384.
XX
PR   14-NOV-2001; 2001JP-00348384.
XX
PA   (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR   WPI; 2003-818165/77.
XX
PT   New promoter DNA derived from Rhodococcus erythropolis KA2-5-1 (FERM P-
PT   16277), useful for efficient production of specific protein.
XX
PS   Example 1; SEQ ID NO 2; 33pp; Japanese.
XX
CC   The invention relates to a promoter DNA (I) which has a fully defined
CC   sequence (S1) of 355 nucleotides, given in the specification or which has
CC   a sequence that hybridises to complementary nucleotides of (S1) and has
CC   promoter activity. The method of the invention is efficient in production
CC   of specific protein using microorganisms. The promoter is expressed
CC   constantly and has high activity. The current sequence represents a
CC   primer related to the novel promoter sequence of the invention.
XX
SQ   Sequence 41 BP; 10 A; 14 C; 5 G; 12 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 19;  DB 10;  Length 41;
  Best Local Similarity  100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          8 CTGTCTCTTATACACATCT 26


RESULT 21
ADG20413
ID   ADG20413 standard; DNA; 41 BP.
XX
AC   ADG20413;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Pseudomonas aeruginosa hcuABC operon primer #1.
XX
KW   intracellular uptake; hydrophobic; hcuABC; operon; desulfurization;
KW   D-light oil; primer; ss.
XX
OS   Pseudomonas aeruginosa.
XX
PN   JP2003219887-A.
XX
PD   05-AUG-2003.
XX
```

PF 30-JAN-2002; 2002JP-00021931.
XX
PR 30-JAN-2002; 2002JP-00021931.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2003-819986/77.
XX
PT Novel gene derived from Pseudomonas aeruginosa NCIMB9571 strain,
PT associated with intracellular uptake of hydrophobic substance useful for
PT transformation of hydrophobic compound.
XX
PS Example 1; SEQ ID NO 5; 23pp; Japanese.
XX
CC The invention relates to gene encoding 3 proteins involved in the
CC intracellular uptake of hydrophobic substances. The operon is named the
CC hcuABC operon. The gene is associated with intracellular hydrophobic
CC substance transport. The proteins and cells transformed with the gene are
CC useful for the desulfurization of D-light oil. This sequence corresponds
CC to a PCR primer used to clone the hcuABC operon sequence.
XX
SQ Sequence 41 BP; 10 A; 14 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 8 CTGTCTCTTATACACATCT 26

RESULT 22
ADF72744
ID ADF72744 standard; DNA; 41 BP.
XX
AC ADF72744;
XX
DT 26-FEB-2004 (first entry)
XX
DE pMOD1 transposon PCR primer SEQ 2 used in promoter search.
XX
KW Desulphurisation; sulphur-containing heterocyclic compound;
KW benzothiophene; dibenzothiophene; alkylated derivative;
KW recombinant bacterium; desulphurisation enzyme;
KW Rhodococcus erythropolis strain KA2-5-1; dszABCD; KAP1 promoter;
KW fossil fuel oil; petroleum; promoter search; transposon; plasmid pMOD1;
KW PCR; primer; ss.
XX
OS Synthetic.
XX
PN JP2003144167-A.
XX
PD 20-MAY-2003.
XX
PF 14-NOV-2001; 2001JP-00348385.
XX
PR 14-NOV-2001; 2001JP-00348385.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2004-075053/08.
XX
PT Desulfurizing sulfur containing heterocyclic compounds involves the use
PT of recombinant microorganisms which contain desulfurization enzyme gene
PT in downstream of a promoter.
XX
PS Claim 1; SEQ ID NO 2; 16pp; Japanese.
XX
CC The invention relates to a method of desulphurising sulphur-containing
CC heterocyclic compounds using recombinant bacteria which contain a gene
CC encoding a desulphurisation enzyme operably linked to a Rhodococcus
CC erythropolis strain KA2-5-1 promoter designated KAP1 (ADF72743). The
CC invention also relates to a recombinant strain of Rhodococcus
CC erythropolis designated MC0203 (FERM P-18595) comprising a
CC desulphurisation enzyme under the control of the KAP1 promoter. The
CC desulphurisation gene used in recombinant microorganisms of the invention
CC selectively cleaves the C-S bond of a sulphur-containing heterocyclic
CC compound, and can be a Sphingomonas sp. strain AD109 decomposition enzyme
CC or a desulphurisation enzyme from Rhodococcus sp. strain IGTS8,
CC Rhodococcus erythropolis strain KA2-5-1 (especially the dszABCD gene
CC (ADF72748)), or Paenibacillus sp. strains A11-1 or A11-2. The method of
CC the invention is useful for desulphurising sulphur-containing
CC heterocyclic compounds such as benzothiophene, dibenzothiophenes and
CC their derivatives, especially their alkylated derivatives. The method
CC permits effective desulphurisation of sulphur-containing compounds in
CC fossil fuel oils such as petroleum under moderate conditions. The present
CC sequence is related to the invention.
XX
SQ Sequence 41 BP; 10 A; 14 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 41;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 8 CTGTCTCTTATACACATCT 26

RESULT 23
ADM35526
ID ADM35526 standard; DNA; 41 BP.
XX
AC ADM35526;
XX
DT 03-JUN-2004 (first entry)
XX
DE Kap1 promoter related transposon primer, SEQ ID No 2.
XX
KW recombinant microorganism; desulfurising; C-S bond;
KW sulfur-containing heterocyclic compound; benzothiophene;
KW dibenzothiophene; alkylated benzothiophene; dibenzothiophene;
KW desulfurisation; sulfate ion; hydroxy compound; kap1 promoter; ss;
KW primer.
XX
OS Rhodococcus erythropolis.
OS Synthetic.
XX
PN JP2004049116-A.
XX
PD 19-FEB-2004.
XX
PF 19-JUL-2002; 2002JP-00211292.
XX
PR 19-JUL-2002; 2002JP-00211292.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2004-174117/17.
XX
PT New recombinant microorganisms belonging to the Mycobacterium genus,
PT useful for continuous desulfurization of sulfur-containing compounds such
PT as (alkylated) benzothiophenes or dibenzothiophenes.
XX
PS Example 1; SEQ ID NO 2; 27pp; Japanese.
XX
CC The invention relates to novel recombinant microorganisms belonging to
CC the Mycobacterium genus comprising an introduced desulfurising gene
CC encoding an enzyme which selectively cleaves the C-S bond, and an
CC introduced promoter, which is constitutively expressed without being
CC suppressed by sulfate ions. The recombinant microorganisms are useful for
CC desulfurising sulfur-containing heterocyclic compounds, which involves
CC culturing the recombinant microorganisms with sulfur-containing compounds
CC such as benzothiophene, dibenzothiophene or their derivatives; or
CC alkylated benzothiophenes or alkylated dibenzothiophenes. The recombinant

```
CC   microorganisms enable continuous desulfurisation of a sulfur-containing
CC   compound without inhibition of the desulfurisation reaction by sulfate
CC   ions and hydroxy compounds. This polynucleotide sequence represents a
CC   primer used in the exemplification of the invention.
XX
SQ   Sequence 41 BP; 10 A; 14 C; 5 G; 12 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 12;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          8 CTGTCTCTTATACACATCT 26


RESULT 24
ADR44424
ID   ADR44424 standard; DNA; 41 BP.
XX
AC   ADR44424;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Plasmid pMOD1-specific PCR primer #1.
XX
KW   sulphur-containing heterocyclic compound; desulphurising gene; dszABCD;
KW   kap1 promoter; fossil fuel oil; pMOD1; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   JP2004242511-A.
XX
PD   02-SEP-2004.
XX
PF   10-FEB-2003; 2003JP-00032686.
XX
PR   10-FEB-2003; 2003JP-00032686.
XX
PA   (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR   WPI; 2004-629584/61.
XX
PT   Desulfurizing sulfur-containing heterocyclic compound using recombinant
PT   microorganisms introduced with desulfurizing gene which encodes
PT   desulfurase and has promoter that does not receive feedback inhibition of
PT   sulfate ion.
XX
PS   Example 1; SEQ ID NO 2; 32pp; Japanese.
XX
CC   The invention comprises a method for desulphurising a sulphur-containing
CC   heterocyclic compound. The method involves the use of recombinant
CC   microorganisms which have been transformed with a desulphurising gene
CC   (e.g. dszABCD). The desulphurising gene is under the control of a
CC   promoter (e.g. kap1) which expresses constantly without suppression by
CC   sulphate ions in the microorganism. The method of the invention is useful
CC   for desulphurising sulphur-containing heterocyclic compounds, such as:
CC   benzothiophene, dibenzothiophene, benzo tetrahydro naphtha thiophene,
CC   benzo naphtha thiophenes, and alkylated benzothiophenes. The method of
CC   the invention is also useful for desulphurising fossil fuel oil
CC   containing sulphur-containing heterocyclic compounds. The present DNA
CC   sequence represents a PCR primer that was used to amplify a region of the
CC   pMOD1 plasmid as part of the preparation of a transposon for a promoter
CC   search.
XX
SQ   Sequence 41 BP; 10 A; 14 C; 5 G; 12 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 13;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          8 CTGTCTCTTATACACATCT 26


RESULT 25
ADC53954
ID   ADC53954 standard; DNA; 42 BP.
XX
AC   ADC53954;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Rhodococcus promoter related primer #SEQ ID 3.
XX
KW   Promoter; protein production; bacterial; PCR; primer; ss.
XX
OS   Rhodococcus erythropolis.
XX
PN   JP2003144166-A.
XX
PD   20-MAY-2003.
XX
PF   14-NOV-2001; 2001JP-00348384.
XX
PR   14-NOV-2001; 2001JP-00348384.
XX
PA   (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR   WPI; 2003-818165/77.
XX
PT   New promoter DNA derived from Rhodococcus erythropolis KA2-5-1 (FERM P-
PT   16277), useful for efficient production of specific protein.
XX
PS   Example 1; SEQ ID NO 3; 33pp; Japanese.
XX
CC   The invention relates to a promoter DNA (I) which has a fully defined
CC   sequence (S1) of 355 nucleotides, given in the specification or which has
CC   a sequence that hybridises to complementary nucleotides of (S1) and has
CC   promoter activity. The method of the invention is efficient in production
CC   of specific protein using microorganisms. The promoter is expressed
CC   constantly and has high activity. The current sequence represents a
CC   primer related to the novel promoter sequence of the invention.
XX
SQ   Sequence 42 BP; 10 A; 13 C; 6 G; 13 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          9 CTGTCTCTTATACACATCT 27


RESULT 26
ADG20414
ID   ADG20414 standard; DNA; 42 BP.
XX
AC   ADG20414;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Pseudomonas aeruginosa hcuABC operon primer #2.
XX
KW   intracellular uptake; hydrophobic; hcuABC; operon; desulfurization;
KW   D-light oil; primer; ss.
XX
OS   Pseudomonas aeruginosa.
XX
PN   JP2003219887-A.
XX
PD   05-AUG-2003.
XX
PF   30-JAN-2002; 2002JP-00021931.
```

XX
PR 30-JAN-2002; 2002JP-00021931.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2003-819986/77.
XX
PT Novel gene derived from Pseudomonas aeruginosa NCIMB9571 strain,
PT associated with intracellular uptake of hydrophobic substance useful for
PT transformation of hydrophobic compound.
XX
PS Example 1; SEQ ID NO 6; 23pp; Japanese.
XX
CC The invention relates to gene encoding 3 proteins involved in the
CC intracellular uptake of hydrophobic substances. The operon is named the
CC hcuABC operon. The gene is associated with intracellular hydrophobic
CC substance transport. The proteins and cells transformed with the gene are
CC useful for the desulfurization of D-light oil. This sequence corresponds
CC to a PCR primer used to clone the hcuABC operon sequence.
XX
SQ Sequence 42 BP; 10 A; 13 C; 6 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 42;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 9 CTGTCTCTTATACACATCT 27

RESULT 27
ADF72745
ID ADF72745 standard; DNA; 42 BP.
XX
AC ADF72745;
XX
DT 26-FEB-2004 (first entry)
XX
DE pMOD1 transposon PCR primer SEQ 3 used in promoter search.
XX
KW Desulphurisation; sulphur-containing heterocyclic compound;
KW benzothiophene; dibenzothiophene; alkylated derivative;
KW recombinant bacterium; desulphurisation enzyme;
KW Rhodococcus erythropolis strain KA2-5-1; dszABCD; KAP1 promoter;
KW fossil fuel oil; petroleum; promoter search; transposon; plasmid pMOD1;
KW PCR; primer; ss.
XX
OS Synthetic.
XX
PN JP2003144167-A.
XX
PD 20-MAY-2003.
XX
PF 14-NOV-2001; 2001JP-00348385.
XX
PR 14-NOV-2001; 2001JP-00348385.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2004-075053/08.
XX
PT Desulfurizing sulfur containing heterocyclic compounds involves the use
PT of recombinant microorganisms which contain desulfurization enzyme gene
PT in downstream of a promoter.
XX
PS Claim 1; SEQ ID NO 3; 16pp; Japanese.
XX
CC The invention relates to a method of desulphurising sulphur-containing
CC heterocyclic compounds using recombinant bacteria which contain a gene
CC encoding a desulphurisation enzyme operably linked to a Rhodococcus
CC erythropolis strain KA2-5-1 promoter designated KAP1 (ADF72743). The
CC invention also relates to a recombinant strain of Rhodococcus
CC erythropolis designated MC0203 (FERM P-18595) comprising a
CC desulphurisation enzyme under the control of the KAP1 promoter. The
CC desulphurisation gene used in recombinant microorganisms of the invention
CC selectively cleaves the C-S bond of a sulphur-containing heterocyclic
CC compound, and can be a Sphingomonas sp. strain AD109 decomposition enzyme
CC or a desulphurisation enzyme from Rhodococcus sp. strain IGTS8,
CC Rhodococcus erythropolis strain KA2-5-1 (especially the dszABCD gene
CC (ADF72748)), or Paenibacillus sp. strains A11-1 or A11-2. The method of
CC the invention is useful for desulphurising sulphur-containing
CC heterocyclic compounds such as benzothiophene, dibenzothiophenes and
CC their derivatives, especially their alkylated derivatives. The method
CC permits effective desulphurisation of sulphur-containing compounds in
CC fossil fuel oils such as petroleum under moderate conditions. The present
CC sequence is related to the invention.
XX
SQ Sequence 42 BP; 10 A; 13 C; 6 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 42;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 9 CTGTCTCTTATACACATCT 27

RESULT 28
ADM35527
ID ADM35527 standard; DNA; 42 BP.
XX
AC ADM35527;
XX
DT 03-JUN-2004 (first entry)
XX
DE Kap1 promoter related transposon primer, SEQ ID No 3.
XX
KW recombinant microorganism; desulfurising; C-S bond;
KW sulfur-containing heterocyclic compound; benzothiophene;
KW dibenzothiophene; alkylated benzothiophene; dibenzothiophene;
KW desulfurisation; sulfate ion; hydroxy compound; kap1 promoter; ss;
KW primer.
XX
OS Rhodococcus erythropolis.
OS Synthetic.
XX
PN JP2004049116-A.
XX
PD 19-FEB-2004.
XX
PF 19-JUL-2002; 2002JP-00211292.
XX
PR 19-JUL-2002; 2002JP-00211292.
XX
PA (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR WPI; 2004-174117/17.
XX
PT New recombinant microorganisms belonging to the Mycobacterium genus,
PT useful for continuous desulfurization of sulfur-containing compounds such
PT as (alkylated) benzothiophenes or dibenzothiophenes.
XX
PS Example 1; SEQ ID NO 3; 27pp; Japanese.
XX
CC The invention relates to novel recombinant microorganisms belonging to
CC the Mycobacterium genus comprising an introduced desulfurising gene
CC encoding an enzyme which selectively cleaves the C-S bond, and an
CC introduced promoter, which is constitutively expressed without being
CC suppressed by sulfate ions. The recombinant microorganisms are useful for
CC desulfurising sulfur-containing heterocyclic compounds, which involves
CC culturing the recombinant microorganisms with sulfur-containing compounds
CC such as benzothiophene, dibenzothiophene or their derivatives; or
CC alkylated benzothiophenes or alkylated dibenzothiophenes. The recombinant
CC microorganisms enable continuous desulfurisation of a sulfur-containing

```
CC   compound without inhibition of the desulfurisation reaction by sulfate
CC   ions and hydroxy compounds. This polynucleotide sequence represents a
CC   primer used in the exemplification of the invention.
XX
SQ   Sequence 42 BP; 10 A; 13 C; 6 G; 13 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 12;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          9 CTGTCTCTTATACACATCT 27


RESULT 29
ADR44425
ID   ADR44425 standard; DNA; 42 BP.
XX
AC   ADR44425;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Plasmid pMOD1-specific PCR primer #2.
XX
KW   sulphur-containing heterocyclic compound; desulphurising gene; dszABCD;
KW   kap1 promoter; fossil fuel oil; pMOD1; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   JP2004242511-A.
XX
PD   02-SEP-2004.
XX
PF   10-FEB-2003; 2003JP-00032686.
XX
PR   10-FEB-2003; 2003JP-00032686.
XX
PA   (KOKU-) ZH KOKUSAI SEKIYU KORYU CENT.
XX
DR   WPI; 2004-629584/61.
XX
PT   Desulfurizing sulfur-containing heterocyclic compound using recombinant
PT   microorganisms introduced with desulfurizing gene which encodes
PT   desulfurase and has promoter that does not receive feedback inhibition of
PT   sulfate ion.
XX
PS   Example 1; SEQ ID NO 3; 32pp; Japanese.
XX
CC   The invention comprises a method for desulphurising a sulphur-containing
CC   heterocyclic compound. The method involves the use of recombinant
CC   microorganisms which have been transformed with a desulphurising gene
CC   (e.g. dszABCD). The desulphurising gene is under the control of a
CC   promoter (e.g. kap1) which expresses constantly without suppression by
CC   sulphate ions in the microorganism. The method of the invention is useful
CC   for desulphurising sulphur-containing heterocyclic compounds, such as:
CC   benzothiophene, dibenzothiophene, benzo tetrahydro naphtha thiophene,
CC   benzo naphtha thiophenes, and alkylated benzothiophenes. The method of
CC   the invention is also useful for desulphurising fossil fuel oil
CC   containing sulphur-containing heterocyclic compounds. The present DNA
CC   sequence represents a PCR primer that was used to amplify a region of the
CC   pMOD1 plasmid as part of the preparation of a transposon for a promoter
CC   search.
XX
SQ   Sequence 42 BP; 10 A; 13 C; 6 G; 13 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 13;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          9 CTGTCTCTTATACACATCT 27


RESULT 30
ACH00838
ID   ACH00838 standard; DNA; 80 BP.
XX
AC   ACH00838;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Primer 98 used in construction of Gene Kelly transposon.
XX
KW   Transposon; gene kelly transposon; RNA polymerase recognition site;
KW   homing endonuclease recognition site; essential gene identification;
KW   antibacterial; antiparasitic; fungicide; pesticide; herbicide; screening;
KW   immunostimulant; PCR; primer; ss.
XX
OS   Synthetic.
XX
PN   WO2003074700-A2.
XX
PD   12-SEP-2003.
XX
PF   05-MAR-2003; 2003WO-GB000918.
XX
PR   05-MAR-2002; 2002GB-00005143.
PR   26-SEP-2002; 2002GB-00022378.
XX
PA   (ARRO-) ARROW THERAPEUTICS LTD.
XX
PI   Maskell DJ,  Charles IG,  Allen A,  Owen P;
XX
DR   WPI; 2003-712894/67.
XX
PT   New transposon comprising an RNA polymerase recognition site and a homing
PT   endonuclease recognition site, useful for identifying genes identify
PT   inhibitors for treating bacterial, fungal or eukaryotic parasite
PT   infection.
XX
PS   Example 1; Page 81; 83pp; English.
XX
CC   The present invention relates to an artificial transposon comprising an
CC   RNA polymerase recognition site and a homing endonuclease recognition
CC   site. The transposon is useful for identifying an essential or a
CC   conditional essential gene. The essential and conditional essential genes
CC   are useful for identifying an inhibitor of transcription and/or
CC   translation of that gene and/or activity of a polypeptide encoded by that
CC   gene. The inhibitor is useful for treating bacterial, fungal or
CC   eukaryotic parasite infection. The bacterium is useful for vaccinating a
CC   human or animal. The present sequence is a PCR primer used to construct
CC   an artificial transposon in the exemplification of the invention
XX
SQ   Sequence 80 BP; 25 A; 15 C; 14 G; 26 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          4 CTGTCTCTTATACACATCT 22


RESULT 31
ADH48035
ID   ADH48035 standard; DNA; 84 BP.
XX
AC   ADH48035;
XX
DT   15-APR-2004  (first entry)
XX
DE   Transposon pPAO16 left arm nucleotide sequence SEQ ID NO:7.
XX
```

KW library analysis; polynucleotide library; cloning vector; characteristic;
KW phenotype; environmental; DNA fragment; molecular characteristic;
KW genetic diversity; metagenomic library; identification; metabolite; drug;
KW enzyme; antibiotic; functional analysis; transposon; gene; ds.
XX
OS Synthetic.
XX
PN EP1386966-A1.
XX
PD 04-FEB-2004.
XX
PF 24-JUL-2002; 2002EP-00291871.
XX
PR 24-JUL-2002; 2002EP-00291871.
XX
PA (LIBR-) LIBRAGEN.
XX
PI Nalin R, Pujic P, Tuphile K, Gillet B;
XX
DR WPI; 2004-124995/13.
XX
PT Analyzing a library of polynucleotides contained in cloning vectors,
PT useful in identifying new metabolites or drugs, comprises modifying the
PT cloning vectors to allow transfer into a host cell for expressing the
PT polynucleotide.
XX
PS Disclosure; SEQ ID NO 7; 88pp; English.
XX
CC The present invention describes a method for analysing a library of
CC polynucleotides contained in cloning vectors having a particular host
CC range. The method comprises: (a) selecting cloning vectors in a library
CC containing a polynucleotide having a particular characteristic; (b)
CC modifying the selected cloning vectors to allow a transfer of the vectors
CC and/or expression of the polynucleotide which they contain into a
CC selected host cell; and (c) analysing the polynucleotide contained in the
CC modified vectors upon transfer of the vectors into the selected host
CC cell. Also described: (1) identifying or cloning polynucleotides encoding
CC a selected phenotype; (2) a transposable nucleic acid construct
CC comprising an origin of transfer flanked by 2 inverted repeats; (3) a
CC library of polynucleotides comprising several environmental DNA fragments
CC cloned into the cloning vectors, where the DNA fragments contain a common
CC molecular characteristic and the cloning vectors are E. coli cloning
CC vectors comprising a target polynucleotide construct allowing transfer or
CC expression of the environmental DNA into a selected host cell distinct
CC from E. coli; (4) a polynucleotide comprising all or a part of a sequence
CC comprising 37500 or 37507 bp (SEQ ID NO: 1 and 2 (ADH48029and ADH48030);
CC and (5) an oligonucleotide comprising the sequence: (i) 5'-
CC GGSCCSKCSSTSDCSRTSGAYACSGC-3' (SEQ ID NO: 3, ADH48031); or (ii) 5'-
CC GCBBSSRYYTCDATSGGRTCSCC-3' (SEQ ID NO: 4, ADH48032). The method is useful
CC for producing and analysing genetic diversity (metagenomic libraries),
CC and to identify and isolate new metabolites, drugs, enzymes or
CC antibiotics. The method has the advantage of high efficient cloning in E.
CC coli and to modify the properties of metagenomic libraries, to allow
CC functional analysis of particular selected clones in any appropriate
CC system, thus making possible the analysis of the a huge diversity of
CC metagenomic libraries. The present sequence represents a transposon
CC pPAO16 left arm nucleotide sequence, which is used in the exemplification
CC of the present invention.
XX
SQ Sequence 84 BP; 17 A; 24 C; 15 G; 28 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 84;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 1 CTGTCTCTTATACACATCT 19

RESULT 32
ACH00837
ID ACH00837 standard; DNA; 85 BP.
XX
AC ACH00837;
XX
DT 12-FEB-2004 (first entry)
XX
DE Primer 97 used in construction of Gene Kelly transposon.
XX
KW Transposon; gene kelly transposon; RNA polymerase recognition site;
KW homing endonuclease recognition site; essential gene identification;
KW antibacterial; antiparasitic; fungicide; pesticide; herbicide; screening;
KW immunostimulant; PCR; primer; ss.
XX
OS Synthetic.
XX
PN WO2003074700-A2.
XX
PD 12-SEP-2003.
XX
PF 05-MAR-2003; 2003WO-GB000918.
XX
PR 05-MAR-2002; 2002GB-00005143.
PR 26-SEP-2002; 2002GB-00022378.
XX
PA (ARRO-) ARROW THERAPEUTICS LTD.
XX
PI Maskell DJ, Charles IG, Allen A, Owen P;
XX
DR WPI; 2003-712894/67.
XX
PT New transposon comprising an RNA polymerase recognition site and a homing
PT endonuclease recognition site, useful for identifying genes identify
PT inhibitors for treating bacterial, fungal or eukaryotic parasite
PT infection.
XX
PS Example 1; Page 81; 83pp; English.
XX
CC The present invention relates to an artificial transposon comprising an
CC RNA polymerase recognition site and a homing endonuclease recognition
CC site. The transposon is useful for identifying an essential or a
CC conditional essential gene. The essential and conditional essential genes
CC are useful for identifying an inhibitor of transcription and/or
CC translation of that gene and/or activity of a polypeptide encoded by that
CC gene. The inhibitor is useful for treating bacterial, fungal or
CC eukaryotic parasite infection. The bacterium is useful for vaccinating a
CC human or animal. The present sequence is a PCR primer used to construct
CC an artificial transposon in the exemplification of the invention
XX
SQ Sequence 85 BP; 22 A; 23 C; 14 G; 26 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||
Db 4 CTGTCTCTTATACACATCT 22

RESULT 33
ADH48036/c
ID ADH48036 standard; DNA; 94 BP.
XX
AC ADH48036;
XX
DT 15-APR-2004 (first entry)
XX
DE Transposon pPAO16 right arm nucleotide sequence SEQ ID NO:8.
XX
KW library analysis; polynucleotide library; cloning vector; characteristic;
KW phenotype; environmental; DNA fragment; molecular characteristic;
KW genetic diversity; metagenomic library; identification; metabolite; drug;
KW enzyme; antibiotic; functional analysis; transposon; gene; ds.
XX

OS Synthetic.
XX
PN EP1386966-A1.
XX
PD 04-FEB-2004.
XX
PF 24-JUL-2002; 2002EP-00291871.
XX
PR 24-JUL-2002; 2002EP-00291871.
XX
PA (LIBR-) LIBRAGEN.
XX
PI Nalin R, Pujic P, Tuphile K, Gillet B;
XX
DR WPI; 2004-124995/13.
XX
PT Analyzing a library of polynucleotides contained in cloning vectors,
PT useful in identifying new metabolites or drugs, comprises modifying the
PT cloning vectors to allow transfer into a host cell for expressing the
PT polynucleotide.
XX
PS Disclosure; SEQ ID NO 8; 88pp; English.
XX
CC The present invention describes a method for analysing a library of
CC polynucleotides contained in cloning vectors having a particular host
CC range. The method comprises: (a) selecting cloning vectors in a library
CC containing a polynucleotide having a particular characteristic; (b)
CC modifying the selected cloning vectors to allow a transfer of the vectors
CC and/or expression of the polynucleotide which they contain into a
CC selected host cell; and (c) analysing the polynucleotide contained in the
CC modified vectors upon transfer of the vectors into the selected host
CC cell. Also described: (1) identifying or cloning polynucleotides encoding
CC a selected phenotype; (2) a transposable nucleic acid construct
CC comprising an origin of transfer flanked by 2 inverted repeats; (3) a
CC library of polynucleotides comprising several environmental DNA fragments
CC cloned into the cloning vectors, where the DNA fragments contain a common
CC molecular characteristic and the cloning vectors are E. coli cloning
CC vectors comprising a target polynucleotide construct allowing transfer or
CC expression of the environmental DNA into a selected host cell distinct
CC from E. coli; (4) a polynucleotide comprising all or a part of a sequence
CC comprising 37500 or 37507 bp (SEQ ID NO: 1 and 2 (ADH48029and ADH48030);
CC and (5) an oligonucleotide comprising the sequence: (i) 5'-
CC GGSCCSKCSSTSDCSRTSGAYACSGC-3' (SEQ ID NO: 3, ADH48031); or (ii) 5'-
CC GCBBSSRYYTCDATSGGRTCSCC-3' (SEQ ID NO: 4, ADH48032). The method is useful
CC for producing and analysing genetic diversity (metagenomic libraries),
CC and to identify and isolate new metabolites, drugs, enzymes or
CC antibiotics. The method has the advantage of high efficient cloning in E.
CC coli and to modify the properties of metagenomic libraries, to allow
CC functional analysis of particular selected clones in any appropriate
CC system, thus making possible the analysis of the a huge diversity of
CC metagenomic libraries. The present sequence represents a transposon
CC pPAO16 right arm nucleotide sequence, which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 94 BP; 28 A; 23 C; 25 G; 18 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||||
Db 94 CTGTCTCTTATACACATCT 76

RESULT 34
AAD58813
ID AAD58813 standard; DNA; 136 BP.
XX
AC AAD58813;
XX
DT 18-DEC-2003 (first entry)
XX
DE Transposon insertion site from clone PP5-8.
XX
KW Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW chimeric; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT misc_feature 1. .30
FT /*tag= a
FT /note= "pMIR86 plasmid DNA fragment"
FT misc_feature 14. .30
FT /*tag= e
FT /note= "9bp duplication of vector sequence at the
FT insertion site"
FT misc_feature 31. .106
FT /*tag= c
FT /note= "pMIR3 plasmid DNA fragment"
FT misc_feature 88. .106
FT /*tag= d
FT /note= "Tn5 mosaic element (ME)"
FT misc_feature 107. .136
FT /*tag= f
FT /note= "pMIR86 plasmid DNA fragment"
FT misc_feature 107. .115
FT /*tag= e
FT /note= "9bp duplication of vector sequence at the
FT insertion site"
XX
PN US2003143740-A1.
XX
PD 31-JUL-2003.
XX
PF 15-OCT-2002; 2002US-00272552.
XX
PR 15-OCT-2001; 2001US-0329474P.
PR 08-NOV-2001; 2001US-0344865P.
XX
PA (WOOD/) WOODDELL C.
PA (HERW/) HERWEIJER H.
PA (WOLF/) WOLFF J A.
XX
PI Wooddell C, Herweijer H, Wolff JA;
XX
DR WPI; 2003-645713/61.
XX
PT Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT comprises delivering a complex between nucleic acid containing a
PT transposon and a transposase specific for the transposon.
XX
PS Example; Page 9; 20pp; English.
XX
CC The invention relates to a method of integrating nucleic acid into the
CC genome of mammalian cells. The method involves forming an integrator
CC complex between the nucleic acid containing a transposon and a
CC transposase specific for the transposon and delivering the integrator
CC complex to a mammalian cell. The method and composition is useful for
CC integrating nucleic acid into the genome of mammalian cells, especially
CC nucleic acids encoding therapeutic proteins for gene therapy. The
CC transposon may be used to integrate large DNA molecules, up to 10 kb or
CC larger, into the genome of a mammalian cell. The present sequence is a
CC transposon insertion site, used to illustrate the method of the
CC invention. This chimeric sequence consists of transposon Tn5 mosaic
CC element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC transposase encoding plasmid pMIR86
XX
SQ Sequence 136 BP; 33 A; 31 C; 34 G; 38 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 136;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19

```
                  ||||||||||||||||||
Db          31 CTGTCTCTTATACACATCT 49


RESULT 35
AAD58813/c
ID   AAD58813 standard; DNA; 136 BP.
XX
AC   AAD58813;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Transposon insertion site from clone PP5-8.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW   chimeric; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .30
FT                   /*tag= a
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    14. .30
FT                   /*tag= e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
FT   misc_feature    31. .106
FT                   /*tag= c
FT                   /note= "pMIR3 plasmid DNA fragment"
FT   misc_feature    88. .106
FT                   /*tag= d
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    107. .136
FT                   /*tag= f
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    107. .115
FT                   /*tag= e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 9; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is a
CC   transposon insertion site, used to illustrate the method of the
CC   invention. This chimeric sequence consists of transposon Tn5 mosaic
CC   element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC   transposase encoding plasmid pMIR86
XX
SQ   Sequence 136 BP; 33 A; 31 C; 34 G; 38 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 136;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
                ||||||||||||||||||
Db        106 CTGTCTCTTATACACATCT 88


RESULT 36
AAD58812
ID   AAD58812 standard; DNA; 137 BP.
XX
AC   AAD58812;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Transposon insertion site from clone PP5-7.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW   chimeric; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .35
FT                   /*tag= a
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    36. .106
FT                   /*tag= c
FT                   /note= "pMIR3 plasmid DNA fragment"
FT   misc_feature    36. .54
FT                   /*tag= b
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    88. .106
FT                   /*tag= d
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    107. .137
FT                   /*tag= f
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    107. .116
FT                   /*tag= e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
```

```
PS   Example; Page 9; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is a
CC   transposon insertion site, used to illustrate the method of the
CC   invention. This chimeric sequence consists of transposon Tn5 mosaic
CC   element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC   transposase encoding plasmid pMIR86
XX
SQ   Sequence 137 BP; 36 A; 35 C; 38 G; 28 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 137;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         36 CTGTCTCTTATACACATCT 54


RESULT 37
AAD58812/c
ID   AAD58812 standard; DNA; 137 BP.
XX
AC   AAD58812;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Transposon insertion site from clone PP5-7.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW   chimeric; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .35
FT                   /*tag=  a
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    36. .106
FT                   /*tag=  c
FT                   /note= "pMIR3 plasmid DNA fragment"
FT   misc_feature    36. .54
FT                   /*tag=  b
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    88. .106
FT                   /*tag=  d
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    107. .137
FT                   /*tag=  f
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    107. .116
FT                   /*tag=  e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 9; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is a
CC   transposon insertion site, used to illustrate the method of the
CC   invention. This chimeric sequence consists of transposon Tn5 mosaic
CC   element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC   transposase encoding plasmid pMIR86
XX
SQ   Sequence 137 BP; 36 A; 35 C; 38 G; 28 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 10;  Length 137;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db        106 CTGTCTCTTATACACATCT 88


RESULT 38
AAD58811
ID   AAD58811 standard; DNA; 137 BP.
XX
AC   AAD58811;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Transposon insertion site from clone PP5-2.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW   chimeric; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .39
FT                   /*tag=  a
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    40. .108
FT                   /*tag=  c
FT                   /note= "pMIR3 plasmid DNA fragment"
FT   misc_feature    40. .58
FT                   /*tag=  b
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    90. .108
FT                   /*tag=  d
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    109. .137
FT                   /*tag=  f
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    109. .117
FT                   /*tag=  e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
```

```
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 9; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is a
CC   transposon insertion site, used to illustrate the method of the
CC   invention. This chimeric sequence consists of transposon Tn5 mosaic
CC   element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC   transposase encoding plasmid pMIR86
XX
SQ   Sequence 137 BP; 39 A; 24 C; 46 G; 28 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 10;  Length 137;
  Best Local Similarity  100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         40 CTGTCTCTTATACACATCT 58


RESULT 39
AAD58811/c
ID   AAD58811 standard; DNA; 137 BP.
XX
AC   AAD58811;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Transposon insertion site from clone PP5-2.
XX
KW   Therapeutic protein; gene therapy; transposon; mosaic element; ME;
KW   chimeric; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1. .39
FT                   /*tag=  a
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    40. .108
FT                   /*tag=  c
FT                   /note= "pMIR3 plasmid DNA fragment"
FT   misc_feature    40. .58
FT                   /*tag=  b
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    90. .108
FT                   /*tag=  d
FT                   /note= "Tn5 mosaic element (ME)"
FT   misc_feature    109. .137
FT                   /*tag=  f
FT                   /note= "pMIR86 plasmid DNA fragment"
FT   misc_feature    109. .117
FT                   /*tag=  e
FT                   /note= "9bp duplication of vector sequence at the
FT                   insertion site"
XX
PN   US2003143740-A1.
XX
PD   31-JUL-2003.
XX
PF   15-OCT-2002; 2002US-00272552.
XX
PR   15-OCT-2001; 2001US-0329474P.
PR   08-NOV-2001; 2001US-0344865P.
XX
PA   (WOOD/) WOODDELL C.
PA   (HERW/) HERWEIJER H.
PA   (WOLF/) WOLFF J A.
XX
PI   Wooddell C,  Herweijer H,  Wolff JA;
XX
DR   WPI; 2003-645713/61.
XX
PT   Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT   comprises delivering a complex between nucleic acid containing a
PT   transposon and a transposase specific for the transposon.
XX
PS   Example; Page 9; 20pp; English.
XX
CC   The invention relates to a method of integrating nucleic acid into the
CC   genome of mammalian cells. The method involves forming an integrator
CC   complex between the nucleic acid containing a transposon and a
CC   transposase specific for the transposon and delivering the integrator
CC   complex to a mammalian cell. The method and composition is useful for
CC   integrating nucleic acid into the genome of mammalian cells, especially
CC   nucleic acids encoding therapeutic proteins for gene therapy. The
CC   transposon may be used to integrate large DNA molecules, up to 10 kb or
CC   larger, into the genome of a mammalian cell. The present sequence is a
CC   transposon insertion site, used to illustrate the method of the
CC   invention. This chimeric sequence consists of transposon Tn5 mosaic
CC   element (ME) and DNA fragments derived from transposon plasmid pMIR3 and
CC   transposase encoding plasmid pMIR86
XX
SQ   Sequence 137 BP; 39 A; 24 C; 46 G; 28 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 10;  Length 137;
  Best Local Similarity  100.0%;  Pred. No. 15;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db        108 CTGTCTCTTATACACATCT 90


RESULT 40
ABA06312/c
ID   ABA06312 standard; DNA; 160 BP.
XX
AC   ABA06312;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Soy bean SCN/SCS resistance related polynucleotide SEQ ID NO 90.
XX
KW   Soy bean; soybean cyst nematode; soybean sudden death syndrome; SCN/SDS;
KW   transgenic plant; Heterodera glycines; Fusarium solani; ds.
XX
```

OS Glycine max.
XX
PN CA2331674-A1.
XX
PD 28-JUL-2001.
XX
PF 29-JAN-2001; 2001CA-02331674.
XX
PR 28-JAN-2000; 2000US-0178811P.
XX
PA (UYSI-) UNIV SOUTHERN ILLINOIS.
XX
PI Lightfoot DA, Meksem K;
XX
DR WPI; 2001-590306/67.
XX
PT Novel genetic marker associated with soybean cyst nematode or soybean
PT sudden death syndrome resistance in soybeans, used to produce resistant
PT cell lines and plants.
XX
PS Disclosure; Page 183; 247pp; English.
XX
CC The invention relates to genetic markers (ABA06224-ABA06344) associated
CC with soybean cyst nematode/soybean sudden death syndrome (SCN/SDS)
CC resistance in soybeans. The genetic markers provide for methods of
CC detecting SCN/SDS, for development of transgenic plant lines resistant to
CC SCN/SDS, especially the SCN Heterodera glycines but also Fusarium solani
CC and isolation of new genes and polypeptides able to provide resistance to
CC H. glycines and F. solani and substances which regulate the expression of
CC these genes
XX
SQ Sequence 160 BP; 43 A; 37 C; 32 G; 48 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 5; Length 160;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         25 CTGTCTCTTATACACATCT 7
```

Search completed: June 13, 2005, 10:16:43
Job time : 202.5 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:53 ; Search time 782.5 Seconds
(without alignments)
1176.548 Million cell updates/sec

Title: US-10-826-573-5
Perfect score: 19
Sequence: 1 ctgtctcttatacacatct 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		19	100.0	19	6	AR072537	AR072537 Sequence
2		19	100.0	19	6	AR121570	AR121570 Sequence
3		19	100.0	19	6	BD251067	BD251067 Method fo
4		19	100.0	19	6	BD064201	BD064201 System fo
c 5		19	100.0	32	6	AX554973	AX554973 Sequence
c 6		19	100.0	38	6	AX554972	AX554972 Sequence
7		19	100.0	80	6	AX828672	AX828672 Sequence
8		19	100.0	84	6	CQ767419	CQ767419 Sequence
9		19	100.0	84	6	CQ774657	CQ774657 Sequence
10		19	100.0	85	6	AX828671	AX828671 Sequence
c 11		19	100.0	94	6	CQ767420	CQ767420 Sequence
c 12		19	100.0	94	6	CQ774658	CQ774658 Sequence
13		19	100.0	294	1	AY271980S1	AY271980 Bartonell
c 14		19	100.0	310	1	AY271968S2	AY271969 Bartonell
c 15		19	100.0	359	1	AY271970S2	AY271971 Bartonell
16		19	100.0	451	1	AY271978S1	AY271978 Bartonell
c 17		19	100.0	454	1	AY271980S2	AY271981 Bartonell
18		19	100.0	465	1	AY271976S1	AY271976 Bartonell
c 19		19	100.0	465	1	AY271976S2	AY271977 Bartonell
c 20		19	100.0	488	1	AY271982S2	AY271983 Bartonell
21		19	100.0	507	1	AY271970S1	AY271970 Bartonell
c 22		19	100.0	552	1	AY271974S2	AY271975 Bartonell
23		19	100.0	585	1	AY271974S1	AY271974 Bartonell
c 24		19	100.0	639	1	AY271978S2	AY271979 Bartonell
c 25		19	100.0	831	6	BD251602	BD251602 Selection
c 26		19	100.0	831	6	AX028303	AX028303 Sequence
27		19	100.0	930	1	AY271982S1	AY271982 Bartonell
28		19	100.0	959	14	AY571855	AY571855 Vulture h
29		19	100.0	1593	6	BD251601	BD251601 Selection
30		19	100.0	1593	6	AX028302	AX028302 Sequence
31		19	100.0	2044	6	AX828670	AX828670 Sequence
c 32		19	100.0	2044	6	AX828670	AX828670 Sequence
33		19	100.0	3418	6	AR072541	AR072541 Sequence
c 34		19	100.0	3418	6	AR072541	AR072541 Sequence
35		19	100.0	3442	12	SCO566337	AJ566337 Synthetic
c 36		19	100.0	3442	12	SCO566337	AJ566337 Synthetic
37		19	100.0	4636	6	BD251603	BD251603 Selection
38		19	100.0	4636	6	AX028304	AX028304 Sequence
39		19	100.0	4740	12	AY453632	AY453632 Expressio
c 40		19	100.0	4740	12	AY453632	AY453632 Expressio
c 41		19	100.0	5349	6	BD251600	BD251600 Selection
c 42		19	100.0	5349	6	AX028301	AX028301 Sequence
43		19	100.0	5387	6	CQ830697	CQ830697 Sequence
c 44		19	100.0	5387	6	CQ830697	CQ830697 Sequence
45		19	100.0	7727	12	AF424805	AF424805 Transposo

ALIGNMENTS

RESULT 1
AR072537
LOCUS AR072537 19 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 8 from patent US 5948622.
ACCESSION AR072537
VERSION AR072537.1 GI:9999301
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Reznikoff,W.S., Goryshin,I.Yu., York,D.L. and Zhou,H.
TITLE System for in vitro transposition
JOURNAL Patent: US 5948622-A 8 07-SEP-1999;
FEATURES Location/Qualifiers
source 1. .19
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCTCTTATACACATCT 19
||||||||||||||||||||
Db 1 CTGTCTCTTATACACATCT 19

RESULT 2
AR121570
LOCUS AR121570 19 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6159736.
ACCESSION AR121570
VERSION AR121570.1 GI:14105146
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Reznikoff,W.S. and Goryshin,I.Y.
TITLE Method for making insertional mutations using a Tn5 synaptic

```
            complex
  JOURNAL   Patent: US 6159736-A 3 12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 64;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 3
BD251067
LOCUS       BD251067                19 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Method for making insertional mutations.
ACCESSION   BD251067
VERSION     BD251067.1  GI:33060837
KEYWORDS    JP 2002531062-A/3.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Reznikoff,W.S. and Goryshin,I.Y.
  TITLE     Method for making insertional mutations
  JOURNAL   Patent: JP 2002531062-A 3 24-SEP-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT     OS   Artificial Sequence
            PN   JP 2002531062-A/3
            PD   24-SEP-2002
            PF   21-SEP-1999 JP 2000574243
            PR   23-SEP-1998 US    09/159363
            PI   WILLIAM S REZNIKOFF,IGOR Y GORYSHIN
            PC   C12N15/09,C12N9/00,C12N15/01,C12Q1/02//G01N33/15,G01N33/50, PC
              G01N33/566,
            PC   C12N15/00,C12N15/00
            CC   Description of Artificial Sequence: Mosaic
            sequence between OE
            CC             and IE
            CC   sequences
            FH   Key             Location/Qualifiers
            FT   source          1. .19
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 64;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 4
BD064201
LOCUS       BD064201                19 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  System for in vitro transposition using modified TN5 transposase.
ACCESSION   BD064201
VERSION     BD064201.1  GI:22609804
KEYWORDS    JP 2001507565-A/4.
SOURCE      Conus quercinus
  ORGANISM  Conus quercinus
            Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
            Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
            Neogastropoda; Conoidea; Conidae; Conus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Reznikoff,W.S., Goryshin,I.Y. and Zhou,H.
  TITLE     System for in vitro transposition using modified TN5 transposase
  JOURNAL   Patent: JP 2001507565-A 4 12-JUN-2001;
            WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT     PN   JP 2001507565-A/4
            PD   12-JUN-2001
            PF   09-SEP-1997 JP 1998512997
            PR   09-SEP-1996 US    08/814877,02-MAY-1997 US    08/850880 PI
            WILLIAM S REZNIKOFF,IGOR YU GORYSHIN,HONG ZHOU PC
            C12N15/55,C12N9/22,C12N15/90,C12N15/85
            CC   Strandedness: Double;
            CC   Topology: Linear;
            CC   /desc= Tn5 variant outer end
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Conus quercinus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:101313"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 64;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 5
AX554973/c
LOCUS       AX554973                32 bp    DNA     linear   PAT 27-NOV-2002
DEFINITION  Sequence 4 from Patent WO0246444.
ACCESSION   AX554973
VERSION     AX554973.1  GI:25898538
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Goryshin,I.Y., Naumann,T.A. and Reznikoff,W.S.
  TITLE     Double transposition methods for manipulating nucleic acids
  JOURNAL   Patent: WO 0246444-A 4 13-JUN-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="LINKER B (COMPRESSED)"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 32;
  Best Local Similarity  100.0%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         32 CTGTCTCTTATACACATCT 14


RESULT 6
AX554972/c
LOCUS       AX554972                38 bp    DNA     linear   PAT 27-NOV-2002
DEFINITION  Sequence 3 from Patent WO0246444.
ACCESSION   AX554972
VERSION     AX554972.1  GI:25898537
KEYWORDS    .
```

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Goryshin,I.Y., Naumann,T.A. and Reznikoff,W.S.
  TITLE     Double transposition methods for manipulating nucleic acids
  JOURNAL   Patent: WO 0246444-A 3 13-JUN-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="LINKER B (FULL LENGTH)"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         38 CTGTCTCTTATACACATCT 20


RESULT 7
AX828672
LOCUS       AX828672                  80 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 3 from Patent WO03074700.
ACCESSION   AX828672
VERSION     AX828672.1  GI:39838610
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Maskell,D.J., Charles,I.G., Allen,A. and Owen,P.
  TITLE     Transposon
  JOURNAL   Patent: WO 03074700-A 3 12-SEP-2003;
            Arrow Therapeutics Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Primer 98"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db          4 CTGTCTCTTATACACATCT 22


RESULT 8
CQ767419
LOCUS       CQ767419                  84 bp    DNA     linear   PAT 04-MAR-2004
DEFINITION  Sequence 7 from Patent EP1386966.
ACCESSION   CQ767419
VERSION     CQ767419.1  GI:45095545
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Nalin,R., Pujic,P., Tuphile,K. and Gillet,B.
  TITLE     Method for the expression of unknown environmental DNA into adapted
            host cells
  JOURNAL   Patent: EP 1386966-A 7 04-FEB-2004;
            Libragen (FR)
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Left arm of
                     pPAOI6 transposon."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 84;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 9
CQ774657
LOCUS       CQ774657                  84 bp    DNA     linear   PAT 06-MAR-2004
DEFINITION  Sequence 7 from Patent WO2004013327.
ACCESSION   CQ774657
VERSION     CQ774657.1  GI:45237873
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Nalin,R., Pujic,P., Tuphile,K. and Gillet,B.
  TITLE     Method for the expression of unknown environmental dna into adapted
            host cells
  JOURNAL   Patent: WO 2004013327-A 7 12-FEB-2004;
            Libragen (FR)
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Left arm of
                     pPAOI6 transposon."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 84;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 10
AX828671
LOCUS       AX828671                  85 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 2 from Patent WO03074700.
ACCESSION   AX828671
VERSION     AX828671.1  GI:39838609
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Maskell,D.J., Charles,I.G., Allen,A. and Owen,P.
  TITLE     Transposon
  JOURNAL   Patent: WO 03074700-A 2 12-SEP-2003;
            Arrow Therapeutics Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Primer 97"
```

```
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 51;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db          4 CTGTCTCTTATACACATCT 22


RESULT 11
CQ767420/c
LOCUS       CQ767420                  94 bp    DNA     linear   PAT 04-MAR-2004
DEFINITION  Sequence 8 from Patent EP1386966.
ACCESSION   CQ767420
VERSION     CQ767420.1  GI:45095547
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Nalin,R., Pujic,P., Tuphile,K. and Gillet,B.
  TITLE     Method for the expression of unknown environmental DNA into adapted
            host cells
  JOURNAL   Patent: EP 1386966-A 8 04-FEB-2004;
            Libragen (FR)
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Rigth arm of
                     pPAOI6 transposon."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         94 CTGTCTCTTATACACATCT 76


RESULT 12
CQ774658/c
LOCUS       CQ774658                  94 bp    DNA     linear   PAT 06-MAR-2004
DEFINITION  Sequence 8 from Patent WO2004013327.
ACCESSION   CQ774658
VERSION     CQ774658.1  GI:45237874
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Nalin,R., Pujic,P., Tuphile,K. and Gillet,B.
  TITLE     Method for the expression of unknown environmental dna into adapted
            host cells
  JOURNAL   Patent: WO 2004013327-A 8 12-FEB-2004;
            Libragen (FR)
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Rigth arm of
                     pPAOI6 transposon."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db         94 CTGTCTCTTATACACATCT 76


RESULT 13
AY271980S1
LOCUS       AY271980S1                294 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 491 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271980
VERSION     AY271980.1  GI:32140326
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 294)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="491"
     misc_feature    1. .270
                     /note="similar to Rhodopseudomonas palustris hypothetical
                     protein"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 1;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 42;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||||
Db        271 CTGTCTCTTATACACATCT 289


RESULT 14
AY271968S2/c
LOCUS       AY271968S2                310 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 31 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271969
VERSION     AY271969.1  GI:32140309
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 310)
```

```
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .310
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="31"
     repeat_region   order(AY271968.1:287. .303,1. .24)
                     /transposon="Tn903"
     misc_feature    25. .310
                     /note="similar to Yersinia pestis
                     D-serine/D-alanine/glycine transporter"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 310;
  Best Local Similarity   100.0%;  Pred. No. 41;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         24 CTGTCTCTTATACACATCT 6


RESULT 15
AY271970S2/c
LOCUS       AY271970S2               359 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 131 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271971
VERSION     AY271971.1  GI:32140312
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 359)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 359)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .359
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="131"
     repeat_region   order(AY271970.1:471. .507,1. .20)
                     /transposon="Tn903"
     misc_feature    21. .359
                     /note="similar to Agrobacterium tumefaciens outer membrane
                     heme receptor"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 359;
  Best Local Similarity   100.0%;  Pred. No. 40;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         20 CTGTCTCTTATACACATCT 2


RESULT 16
AY271978S1
LOCUS       AY271978S1               451 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 337 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271978
VERSION     AY271978.1  GI:32140323
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 451)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="337"
     misc_feature    1. .427
                     /note="similar to Brucella melitensis hypothetical
                     protein"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db        428 CTGTCTCTTATACACATCT 446


RESULT 17
AY271980S2/c
LOCUS       AY271980S2               454 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 491 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271981
VERSION     AY271981.1  GI:32140327
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 454)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
```

```
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="491"
     repeat_region   order(AY271980.1:271. .294,1. .24)
                     /transposon="Tn903"
     misc_feature    25. .454
                     /note="similar to Rhodopseudomonas palustris hypothetical
                     protein"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 454;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         24 CTGTCTCTTATACACATCT 6


RESULT 18
AY271976S1
LOCUS       AY271976S1              465 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 188 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271976
VERSION     AY271976.1  GI:32140320
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 465)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .465
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="188"
     misc_feature    1. .439
                     /note="similar to Brucella melitensis hypothetical
                     protein"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db        440 CTGTCTCTTATACACATCT 458


RESULT 19
AY271976S2/c
LOCUS       AY271976S2              465 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 188 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271977
VERSION     AY271977.1  GI:32140321
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 465)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 465)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .465
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="188"
     repeat_region   order(AY271976.1:440. .465,1. .26)
                     /transposon="Tn903"
     misc_feature    27. .465
                     /note="similar to Brucella melitensis hypothetical
                     protein"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         26 CTGTCTCTTATACACATCT 8


RESULT 20
AY271982S2/c
LOCUS       AY271982S2              488 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 859 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271983
VERSION     AY271983.1  GI:32140330
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 488)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="859"
     repeat_region   order(AY271982.1:894. .930,1. .19)
                     /transposon="Tn903"
     misc_feature    20. .488
                     /note="similar to Brucella melitensis kinesin-like
                     protein"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         20 CTGTCTCTTATACACATCT 2


RESULT 21
AY271970S1
LOCUS       AY271970S1               507 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 131 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271970
VERSION     AY271970.1  GI:32140311
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 507)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="131"
     misc_feature    1. .470
                     /note="similar to Agrobacterium tumefaciens outer membrane
                     heme receptor"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 507;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db        471 CTGTCTCTTATACACATCT 489


RESULT 22
AY271974S2/c
LOCUS       AY271974S2               552 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 169 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271975
VERSION     AY271975.1  GI:32140318
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 552)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .552
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="169"
     repeat_region   order(AY271974.1:555. .585,1. .20)
                     /transposon="Tn903"
     misc_feature    21. .552
                     /note="similar to Brucella melitensis ATP synthase A chain
                     and ATP synthase subunit Bprime"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 1;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         20 CTGTCTCTTATACACATCT 2


RESULT 23
AY271974S1
LOCUS       AY271974S1               585 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 169 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271974
VERSION     AY271974.1  GI:32140317
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 585)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE.    Direct Submission
```

```
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="169"
     misc_feature    1. .554
                     /note="similar to Brucella melitensis ATP synthase A chain
                     and ATP synthase subunit Bprime"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db        555 CTGTCTCTTATACACATCT 573


RESULT 24
AY271978S2/c
LOCUS       AY271978S2               639 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 337 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271979
VERSION     AY271979.1  GI:32140324
KEYWORDS    .
SEGMENT     2 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 639)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="337"
     repeat_region   order(AY271978.1:428. .451,1. .24)
                     /transposon="Tn903"
     misc_feature    25. .639
                     /note="similar to Brucella melitensis hypothetical
                     protein"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 639;
  Best Local Similarity   100.0%;  Pred. No. 37;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         24 CTGTCTCTTATACACATCT 6


RESULT 25
BD251602/c
LOCUS       BD251602                 831 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Selection of animal based on character imprinted by parent.
ACCESSION   BD251602
VERSION     BD251602.1  GI:33061372
KEYWORDS    JP 2002535963-A/122.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   Andersson,L., Georges,M., Spincemaille,G. and Nezer,C.D.A.
  TITLE     Selection of animal based on character imprinted by parent
  JOURNAL   Patent: JP 2002535963-A 122 29-OCT-2002;
            UNIVERSITY OF LIEGE,MELICA HB,SEGHERS GENTEC NV
COMMENT     OS   Sus scrofa (pig)
            PN   JP 2002535963-A/122
            PD   29-OCT-2002
            PF   16-DEC-1999 JP 2000588390
            PR   16-DEC-1998 EP   98204291.3
            PI   LEIF ANDERSSON,MICHEL GEORGES,GEERT SPINCEMAILLE, PI   CARINE
            DANIELLE ANDREE NEZER
            PC   C12N15/09,A01K67/027,C12N5/06,C12Q1/68,C12N15/00,C12N5/00 CC
            /note='Contig 5, figure 8'
            FH   Key             Location/Qualifiers
            FT   source          1. .831
            FT                   /organism='Sus scrofa (pig)'.
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9823"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 831;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         21 CTGTCTCTTATACACATCT 3


RESULT 26
AX028303/c
LOCUS       AX028303                 831 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 122 from Patent WO0036143.
ACCESSION   AX028303
VERSION     AX028303.1  GI:10189110
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1
  AUTHORS   Georges,M., Spincemaille,G. and Andersson,L.
  TITLE     Selecting animals for parentally imprinted traits
  JOURNAL   Patent: WO 0036143-A 122 22-JUN-2000;
            SEGHERSGENTEC N V (BE) ; GEORGES MICHEL (BE) ; UNIV LIEGE (BE) ;
            SPINCEMAILLE GEERT (BE) ; MELICA HB (SE) ; ANDERSSON LEIF (SE)
FEATURES             Location/Qualifiers
     source          1. .831
                     /organism="Sus scrofa"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9823"
                     /note="Contig 5, figure 8"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 831;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db         21 CTGTCTCTTATACACATCT 3


RESULT 27
AY271982S1
LOCUS       AY271982S1               930 bp    DNA     linear   BCT 29-AUG-2003
DEFINITION  Bartonella henselae clone 859 transposon Tn903-interrupted genomic
            sequence.
ACCESSION   AY271982
VERSION     AY271982.1  GI:32140329
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Bartonella henselae
  ORGANISM  Bartonella henselae
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bartonellaceae; Bartonella.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Rapid and efficient transposon mutagenesis of Bartonella henselae
            by transposome technology
  JOURNAL   Gene 313, 103-109 (2003)
REFERENCE   2  (bases 1 to 930)
  AUTHORS   Riess,T., Anderson,B., Fackelmayer,A., Autenrieth,I.B. and
            Kempf,V.A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2003) Institute for Medical Microbiology,
            University of Tuebingen, Elfriede Aulhorn Str. 6, Tuebingen 72076,
            Germany
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Bartonella henselae"
                     /mol_type="genomic DNA"
                     /strain="Marseille"
                     /db_xref="taxon:38323"
                     /clone="859"
     misc_feature    1. .893
                     /note="similar to Brucella melitensis kinesin-like
                     protein"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 1;  Length 930;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db        894 CTGTCTCTTATACACATCT 912


RESULT 28
AY571855
LOCUS       AY571855                 959 bp    DNA     linear   VRL 01-AUG-2004
DEFINITION  Vulture herpesvirus clone 1 nonfunctional DNA helicase-primase
            component (UL5) gene, partial sequence.
ACCESSION   AY571855
VERSION     AY571855.1  GI:50593442
KEYWORDS    .
SOURCE      Vulture herpesvirus
  ORGANISM  Vulture herpesvirus
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   Cardoso,M.J., Hyatt,A., Selleck,P., Lowther,S., Cunningham,A. and
            Boyle,D.B.
  TITLE     Phylogenetic analysis of the DNA polymerase gene of a novel
            Alphaherpesvirus isolated from an Indian Gyps vulture
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 959)
  AUTHORS   Cardoso,M.J. and Boyle,D.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2004) Australian Animal Health Laboratories,
            CSIRO Livestock Industries, Private Bag 24, Geelong, VIC 3220,
            Australia
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Vulture herpesvirus"
                     /mol_type="genomic DNA"
                     /specific_host="Gyps indicus"
                     /db_xref="taxon:285986"
                     /clone="1"
                     /country="India"
     gene            <1. .>959
                     /gene="UL5"
     misc_feature    <1. .>959
                     /gene="UL5"
                     /note="nonfunctional DNA helicase-primase component due to
                     mutation"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 14;  Length 959;
  Best Local Similarity   100.0%;  Pred. No. 35;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db        933 CTGTCTCTTATACACATCT 951


RESULT 29
BD251601
LOCUS       BD251601                 1593 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Selection of animal based on character imprinted by parent.
ACCESSION   BD251601
VERSION     BD251601.1  GI:33061371
KEYWORDS    JP 2002535963-A/121.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1593)
  AUTHORS   Andersson,L., Georges,M., Spincemaille,G. and Nezer,C.D.A.
  TITLE     Selection of animal based on character imprinted by parent
  JOURNAL   Patent: JP 2002535963-A 121 29-OCT-2002;
            UNIVERSITY OF LIEGE,MELICA HB,SEGHERS GENTEC NV
COMMENT     OS   Sus scrofa (pig)
            PN   JP 2002535963-A/121
            PD   29-OCT-2002
            PF   16-DEC-1999 JP 2000588390
            PR   16-DEC-1998 EP   98204291.3
            PI   LEIF ANDERSSON,MICHEL GEORGES,GEERT SPINCEMAILLE, PI   CARINE
            DANIELLE ANDREE NEZER
            PC   C12N15/09,A01K67/027,C12N5/06,C12Q1/68,C12N15/00,C12N5/00 CC
            /note='Contig 4, figure 8'
            FH   Key             Location/Qualifiers
            FT   source          1. .1593
            FT                   /organism='Sus scrofa (pig)'.
FEATURES             Location/Qualifiers
     source          1. .1593
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9823"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1593;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              ||||||||||||||||||
Db       1567 CTGTCTCTTATACACATCT 1585


RESULT 30
```

```
AX028302
LOCUS       AX028302               1593 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 121 from Patent WO0036143.
ACCESSION   AX028302
VERSION     AX028302.1  GI:10189109
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1
  AUTHORS   Georges,M., Spincemaille,G. and Andersson,L.
  TITLE     Selecting animals for parentally imprinted traits
  JOURNAL   Patent: WO 0036143-A 121 22-JUN-2000;
            SEGHERSGENTEC N V (BE) ; GEORGES MICHEL (BE) ; UNIV LIEGE (BE) ;
            SPINCEMAILLE GEERT (BE) ; MELICA HB (SE) ; ANDERSSON LEIF (SE)
FEATURES             Location/Qualifiers
     source          1. .1593
                     /organism="Sus scrofa"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9823"
                     /note="Contig 4, figure 8"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 1593;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       1567 CTGTCTCTTATACACATCT 1585


RESULT 31
AX828670
LOCUS       AX828670               2044 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 1 from Patent WO03074700.
ACCESSION   AX828670
VERSION     AX828670.1  GI:39838608
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Maskell,D.J., Charles,I.G., Allen,A. and Owen,P.
  TITLE     Transposon
  JOURNAL   Patent: WO 03074700-A 1 12-SEP-2003;
            Arrow Therapeutics Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .2044
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Gene Kelly transposon"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 2044;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 32
AX828670/c
LOCUS       AX828670               2044 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 1 from Patent WO03074700.
ACCESSION   AX828670
VERSION     AX828670.1  GI:39838608
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Maskell,D.J., Charles,I.G., Allen,A. and Owen,P.
  TITLE     Transposon
  JOURNAL   Patent: WO 03074700-A 1 12-SEP-2003;
            Arrow Therapeutics Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .2044
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Gene Kelly transposon"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 2044;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       2044 CTGTCTCTTATACACATCT 2026


RESULT 33
AR072541
LOCUS       AR072541               3418 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 12 from patent US 5948622.
ACCESSION   AR072541
VERSION     AR072541.1  GI:9999305
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3418)
  AUTHORS   Reznikoff,W.S., Goryshin,I.Yu., York,D.L. and Zhou,H.
  TITLE     System for in vitro transposition
  JOURNAL   Patent: US 5948622-A 12 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .3418
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 3418;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       1337 CTGTCTCTTATACACATCT 1355


RESULT 34
AR072541/c
LOCUS       AR072541               3418 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 12 from patent US 5948622.
ACCESSION   AR072541
VERSION     AR072541.1  GI:9999305
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3418)
  AUTHORS   Reznikoff,W.S., Goryshin,I.Yu., York,D.L. and Zhou,H.
  TITLE     System for in vitro transposition
  JOURNAL   Patent: US 5948622-A 12 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .3418
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
  Query Match           100.0%;  Score 19;  DB 6;  Length 3418;
  Best Local Similarity  100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db         80 CTGTCTCTTATACACATCT 62


RESULT 35
SCO566337
LOCUS       SCO566337               3442 bp    DNA     linear   SYN 06-MAY-2004
DEFINITION  Synthetic construct for Streptomyces coelicolor transposon Tn5062.
ACCESSION   AJ566337
VERSION     AJ566337.1  GI:31711419
KEYWORDS    aac(3)IV gene; apramycin resistance; egfp gene; EGFP protein;
            origin of transfer; oriT.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Bishop,A., Fielding,S., Dyson,P. and Herron,P.
  TITLE     Systematic insertional mutagenesis of a streptomycete genome: a
            link between osmoadaptation and antibiotic production
  JOURNAL   Genome Res. 14 (5), 893-900 (2004)
   PUBMED   15078860
REFERENCE   2  (bases 1 to 3442)
  AUTHORS   Herron,P.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-2003) Herron P.R., School of Biological Sciences,
            University of Wales Swansea, Singleton Park, Swansea, Wales, SA2
            8PP, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .3442
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /note="Streptomyces coelicolor transposon Tn5062"
     misc_feature    complement(1. .19)
                     /note="Tn5 mosaic end"
     misc_feature    46. .56
                     /note="three frame translational stop"
     RBS             65. .69
                     /note="Streptomyces consensus ribosome binding site"
     gene            77. .796
                     /gene="egfp"
     CDS             77. .796
                     /gene="egfp"
                     /codon_start=1
                     /transl_table=11
                     /product="EGFP protein"
                     /protein_id="CAD97424.1"
                     /db_xref="GI:31711420"
                     /translation="MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT
                     LKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFK
                     DDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN
                     GIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH
                     MVLLEFVTAAGITLGMDELYK"
     terminator      complement(849. .971)
                     /note="T4 terminator"
     gene            complement(1316. .2101)
                     /gene="aac(3)IV"
     CDS             complement(1316. .2101)
                     /gene="aac(3)IV"
                     /codon_start=1
                     /transl_table=11
                     /product="apramycin resistance"
                     /protein_id="CAD97425.1"
                     /db_xref="GI:31711421"
                     /translation="MQYEWRKAELIGQLLNLGVTPGGVLLVHSSFRSVRPLEDGPLGL
                     IEALRAALGPGGTLVMPSWSGLDDEPFDPATSPVTPDLGVVSDTFWRLPNVKRSAHPF
                     AFAAAGPQAEQIISDPLPLPPHSPASPVARVHELDGQVLLLGVGHDANTTLHLAELMA
                     KVPYGVPRHCTILQDGKLVRVDYLENDHCCERFALADRWLKEKSLQKEGPVGHAFARL
                     IRSRDIVATALGQLGRDPLIFLHPPEGGMRRMRCRSPVDWLSS"
     terminator      2451. .2573
                     /note="T4 terminator"
     oriT            complement(2756. .2856)
                     /note="RK2 origin of transfer"
     misc_feature    3424. .3442
                     /note="Tn5 mosaic end"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 12;  Length 3442;
  Best Local Similarity  100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db          1 CTGTCTCTTATACACATCT 19


RESULT 36
SCO566337/c
LOCUS       SCO566337               3442 bp    DNA     linear   SYN 06-MAY-2004
DEFINITION  Synthetic construct for Streptomyces coelicolor transposon Tn5062.
ACCESSION   AJ566337
VERSION     AJ566337.1  GI:31711419
KEYWORDS    aac(3)IV gene; apramycin resistance; egfp gene; EGFP protein;
            origin of transfer; oriT.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Bishop,A., Fielding,S., Dyson,P. and Herron,P.
  TITLE     Systematic insertional mutagenesis of a streptomycete genome: a
            link between osmoadaptation and antibiotic production
  JOURNAL   Genome Res. 14 (5), 893-900 (2004)
   PUBMED   15078860
REFERENCE   2  (bases 1 to 3442)
  AUTHORS   Herron,P.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-2003) Herron P.R., School of Biological Sciences,
            University of Wales Swansea, Singleton Park, Swansea, Wales, SA2
            8PP, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .3442
                     /organism="synthetic construct"
                     /mol_type="other DNA"
                     /db_xref="taxon:32630"
                     /note="Streptomyces coelicolor transposon Tn5062"
     misc_feature    complement(1. .19)
                     /note="Tn5 mosaic end"
     misc_feature    46. .56
                     /note="three frame translational stop"
     RBS             65. .69
                     /note="Streptomyces consensus ribosome binding site"
     gene            77. .796
                     /gene="egfp"
     CDS             77. .796
                     /gene="egfp"
                     /codon_start=1
                     /transl_table=11
                     /product="EGFP protein"
                     /protein_id="CAD97424.1"
                     /db_xref="GI:31711420"
                     /translation="MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT
                     LKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFK
                     DDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN
                     GIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH
                     MVLLEFVTAAGITLGMDELYK"
     terminator      complement(849. .971)
                     /note="T4 terminator"
     gene            complement(1316. .2101)
                     /gene="aac(3)IV"
     CDS             complement(1316. .2101)
                     /gene="aac(3)IV"
```

```
                     /codon_start=1
                     /transl_table=11
                     /product="apramycin resistance"
                     /protein_id="CAD97425.1"
                     /db_xref="GI:31711421"
                     /translation="MQYEWRKAELIGQLLNLGVTPGGVLLVHSSFRSVRPLEDGPLGL
                     IEALRAALGPGGTLVMPSWSGLDDEPFDPATSPVTPDLGVVSDTFWRLPNVKRSAHPF
                     AFAAAGPQAEQIISDPLPLPPHSPASPVARVHELDGQVLLLGVGHDANTTLHLAELMA
                     KVPYGVPRHCTILQDGKLVRVDYLENDHCCERFALADRWLKEKSLQKEGPVGHAFARL
                     IRSRDIVATALGQLGRDPLIFLHPPEGGMRRMRCRSPVDWLSS"
     terminator      2451. .2573
                     /note="T4 terminator"
     oriT            complement(2756. .2856)
                     /note="RK2 origin of transfer"
     misc_feature    3424. .3442
                     /note="Tn5 mosaic end"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 12;  Length 3442;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       3442 CTGTCTCTTATACACATCT 3424


RESULT 37
BD251603
LOCUS       BD251603                4636 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Selection of animal based on character imprinted by parent.
ACCESSION   BD251603
VERSION     BD251603.1  GI:33061373
KEYWORDS    JP 2002535963-A/123.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 4636)
  AUTHORS   Andersson,L., Georges,M., Spincemaille,G. and Nezer,C.D.A.
  TITLE     Selection of animal based on character imprinted by parent
  JOURNAL   Patent: JP 2002535963-A 123 29-OCT-2002;
            UNIVERSITY OF LIEGE,MELICA HB,SEGHERS GENTEC NV
COMMENT     OS   Sus scrofa (pig)
            PN   JP 2002535963-A/123
            PD   29-OCT-2002
            PF   16-DEC-1999 JP 2000588390
            PR   16-DEC-1998 EP   98204291.3
            PI   LEIF ANDERSSON,MICHEL GEORGES,GEERT SPINCEMAILLE, PI   CARINE
            DANIELLE ANDREE NEZER
            PC   C12N15/09,A01K67/027,C12N5/06,C12Q1/68,C12N15/00,C12N5/00 CC
            /note='Contig 6, figure 8'
            FH   Key             Location/Qualifiers
            FT   source          1. .4636
            FT                   /organism='Sus scrofa (pig)'.
FEATURES             Location/Qualifiers
     source          1. .4636
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9823"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 4636;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       4608 CTGTCTCTTATACACATCT 4626


RESULT 38
AX028304
LOCUS       AX028304                4636 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 123 from Patent WO0036143.
ACCESSION   AX028304
VERSION     AX028304.1  GI:10189111
KEYWORDS    .
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1
  AUTHORS   Georges,M., Spincemaille,G. and Andersson,L.
  TITLE     Selecting animals for parentally imprinted traits
  JOURNAL   Patent: WO 0036143-A 123 22-JUN-2000;
            SEGHERSGENTEC N V (BE) ; GEORGES MICHEL (BE) ; UNIV LIEGE (BE) ;
            SPINCEMAILLE GEERT (BE) ; MELICA HB (SE) ; ANDERSSON LEIF (SE)
FEATURES             Location/Qualifiers
     source          1. .4636
                     /organism="Sus scrofa"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9823"
                     /note="Contig 6, figure 8"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 4636;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       4608 CTGTCTCTTATACACATCT 4626


RESULT 39
AY453632
LOCUS       AY453632                4740 bp    DNA     circular SYN 11-JAN-2004
DEFINITION  Expression vector pEMCJH04, complete sequence.
ACCESSION   AY453632
VERSION     AY453632.1  GI:40646599
KEYWORDS    .
SOURCE      Expression vector pEMCJH04
  ORGANISM  Expression vector pEMCJH04
            other sequences; artificial sequences; vectors.
REFERENCE   1  (bases 1 to 4740)
  AUTHORS   Hays,J.P., Eadie,K., Verduin,C.M., Verbrugh,H. and Van Belkum,A.
  TITLE     A novel plasmid isolated from Moraxella catarrhalis can be used to
            express heterologous proteins within this species
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4740)
  AUTHORS   Hays,J.P., Eadie,K., Verduin,C.M., Verbrugh,H. and Van Belkum,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-OCT-2003) Medical Microbiology and Infectious
            Diseases, Erasmus Mc, Dr. Molewaterplein 40, Rotterdam POSTBUS
            2040, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .4740
                     /organism="Expression vector pEMCJH04"
                     /mol_type="other DNA"
                     /db_xref="taxon:260048"
                     /clone="3.9"
                     /lab_host="Moraxella catarrhalis"
                     /note="derived from plasmid pEMCJH03 containing random
                     insert of transposon EZ::TN<Kan2>"
     oriT            92. .204
                     /note="putative"
     CDS             230. .1147
                     /codon_start=1
                     /transl_table=11
                     /product="putative replicase"
                     /protein_id="AAR88169.1"
                     /db_xref="GI:40646600"
                     /translation="MQNINFSYDLALIQAQILEKLPQKPYCTDDLGYLVVRPKEIAIQ
                     KRYIQHNPPCFTSFLVFDLDTELGAVAWHDADLPLPTWTTQNPDNGHAHIAYALKTPV
                     VTTEQGSNRAIDYLAKIQAGMARKLGADTGYSGLITKNPLHNHWRTTVWNTKPYELGE
```

```
                     LADRVELAPLTPKERELGLGRNCTLFDTVRKWAYIAIREHRGGRYDDWYARVLATVQA
                     ENMAFLEPLPYSEIKATAKSIAVYCWKNDGHCYNEFIYRQRLKGAIGGKKSKRPPMVD
                     SERTLQPWIELGISRRTYYRDLNDNHSQK"
     rep_origin      1053. .1414
                     /note="putative oriV"
     gene            1488. .>1746
                     /gene="putative mobilization gene C"
                     /note="disrupted by transposon insertion"
     repeat_region   1747. .2967
                     /transposon="EZ::TN<Kan2>"
     CDS             1884. .2699
                     /function="kanamycin resistance"
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside 3'-phosphotransferase"
                     /protein_id="AAR88170.1"
                     /db_xref="GI:40646601"
                     /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                     GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                     GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                     VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
                     RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
     gene            <2976. .3058
                     /gene="putative mobilization gene C"
                     /note="disrupted by transposon insertion"
     CDS             3281. .4543
                     /codon_start=1
                     /transl_table=11
                     /product="putative mobilization protein A"
                     /protein_id="AAR88171.1"
                     /db_xref="GI:40646602"
                     /translation="MASFERTLMAGLNQDRYNILWVEHTDKDRLELNFLIPKVDLGTG
                     KAMNPYFDKTDRGLVDVWKQVINYDYGLHDPDDPKNRQTLVTVKDLPKSKQEFKQALT
                     AVLEQKILADEIKDHADIIKELENMGLEIARTTPTAISIKDPDGGRNIRLKGEIYEQT
                     FTANQATERESQRASESYRNELEQRISRVRDELTSRIEAKSAFNATRYKTIPSREQSP
                     NEQAHGIQDPSRGSNGDFVINPDSIRGVQSVLGQENSHTARAIRRYTTGDRQQPPTSQ
                     STNESTGNGTGRQDLHRQQDEQSQNMAKQRQTTNHGATLNVKAIPERVRAIATRARTL
                     LVIARDGKSDAQATDRAITATNSGLRDRKQQANDRKQRTVEIIGVAKNAGAGIDQHHA
                     EQVRLQQQQARQARQQTKDEPKKLGLDR"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 12;  Length 4740;
  Best Local Similarity  100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       1747 CTGTCTCTTATACACATCT 1765


RESULT 40
AY453632/c
LOCUS       AY453632                4740 bp    DNA     circular SYN 11-JAN-2004
DEFINITION  Expression vector pEMCJH04, complete sequence.
ACCESSION   AY453632
VERSION     AY453632.1  GI:40646599
KEYWORDS    .
SOURCE      Expression vector pEMCJH04
  ORGANISM  Expression vector pEMCJH04
            other sequences; artificial sequences; vectors.
REFERENCE   1  (bases 1 to 4740)
  AUTHORS   Hays,J.P., Eadie,K., Verduin,C.M., Verbrugh,H. and Van Belkum,A.
  TITLE     A novel plasmid isolated from Moraxella catarrhalis can be used to
            express heterologous proteins within this species
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4740)
  AUTHORS   Hays,J.P., Eadie,K., Verduin,C.M., Verbrugh,H. and Van Belkum,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-OCT-2003) Medical Microbiology and Infectious
            Diseases, Erasmus Mc, Dr. Molewaterplein 40, Rotterdam POSTBUS
            2040, The Netherlands
FEATURES             Location/Qualifiers
     source          1. .4740
                     /organism="Expression vector pEMCJH04"
                     /mol_type="other DNA"
                     /db_xref="taxon:260048"
                     /clone="3.9"
                     /lab_host="Moraxella catarrhalis"
                     /note="derived from plasmid pEMCJH03 containing random
                     insert of transposon EZ::TN<Kan2>"
     oriT            92. .204
                     /note="putative"
     CDS             230. .1147
                     /codon_start=1
                     /transl_table=11
                     /product="putative replicase"
                     /protein_id="AAR88169.1"
                     /db_xref="GI:40646600"
                     /translation="MQNINFSYDLALIQAQILEKLPQKPYCTDDLGYLVVRPKEIAIQ
                     KRYIQHNPPCFTSFLVFDLDTELGAVAWHDADLPLPTWTTQNPDNGHAHIAYALKTPV
                     VTTEQGSNRAIDYLAKIQAGMARKLGADTGYSGLITKNPLHNHWRTTVWNTKPYELGE
                     LADRVELAPLTPKERELGLGRNCTLFDTVRKWAYIAIREHRGGRYDDWYARVLATVQA
                     ENMAFLEPLPYSEIKATAKSIAVYCWKNDGHCYNEFIYRQRLKGAIGGKKSKRPPMVD
                     SERTLQPWIELGISRRTYYRDLNDNHSQK"
     rep_origin      1053. .1414
                     /note="putative oriV"
     gene            1488. .>1746
                     /gene="putative mobilization gene C"
                     /note="disrupted by transposon insertion"
     repeat_region   1747. .2967
                     /transposon="EZ::TN<Kan2>"
     CDS             1884. .2699
                     /function="kanamycin resistance"
                     /codon_start=1
                     /transl_table=11
                     /product="aminoglycoside 3'-phosphotransferase"
                     /protein_id="AAR88170.1"
                     /db_xref="GI:40646601"
                     /translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY
                     GKPDAPELFLKHGKGSVANDVTDEMVRLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP
                     GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFRLAQAQSRMNNGL
                     VDASDFDDERNGWPVEQVWKEMHKLLPFSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG
                     RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF"
     gene            <2976. .3058
                     /gene="putative mobilization gene C"
                     /note="disrupted by transposon insertion"
     CDS             3281. .4543
                     /codon_start=1
                     /transl_table=11
                     /product="putative mobilization protein A"
                     /protein_id="AAR88171.1"
                     /db_xref="GI:40646602"
                     /translation="MASFERTLMAGLNQDRYNILWVEHTDKDRLELNFLIPKVDLGTG
                     KAMNPYFDKTDRGLVDVWKQVINYDYGLHDPDDPKNRQTLVTVKDLPKSKQEFKQALT
                     AVLEQKILADEIKDHADIIKELENMGLEIARTTPTAISIKDPDGGRNIRLKGEIYEQT
                     FTANQATERESQRASESYRNELEQRISRVRDELTSRIEAKSAFNATRYKTIPSREQSP
                     NEQAHGIQDPSRGSNGDFVINPDSIRGVQSVLGQENSHTARAIRRYTTGDRQQPPTSQ
                     STNESTGNGTGRQDLHRQQDEQSQNMAKQRQTTNHGATLNVKAIPERVRAIATRARTL
                     LVIARDGKSDAQATDRAITATNSGLRDRKQQANDRKQRTVEIIGVAKNAGAGIDQHHA
                     EQVRLQQQQARQARQQTKDEPKKLGLDR"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 12;  Length 4740;
  Best Local Similarity  100.0%;  Pred. No. 27;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGTCTCTTATACACATCT 19
              |||||||||||||||||||
Db       2967 CTGTCTCTTATACACATCT 2949


Search completed: June 13, 2005, 10:42:54
Job time : 783.5 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:53 ; Search time 1597.5 Seconds
(without alignments)
452.721 Million cell updates/sec

Title: US-10-826-573-3
Perfect score: 19
Sequence: 1 ctgactcttatacacaagt 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19	100.0	264	8	AQ215171	AQ215171 1G2 UCBPP
	2	19	100.0	482	2	BE390386	BE390386 601286635
c	3	19	100.0	498	8	AQ215170	AQ215170 8C12 UCBP
	4	19	100.0	503	4	BG338136	BG338136 602435977
	5	19	100.0	507	4	BG340312	BG340312 602438538
	6	19	100.0	509	2	BF685376	BF685376 602142150
	7	19	100.0	509	4	BG339203	BG339203 602437059
	8	19	100.0	518	4	BG759369	BG759369 602711887
	9	19	100.0	535	4	BM010253	BM010253 603631110
	10	19	100.0	539	4	BF971706	BF971706 602239965
	11	19	100.0	539	4	BF974003	BF974003 602243170
	12	19	100.0	539	4	BG755947	BG755947 602716423
	13	19	100.0	539	5	BP033973	BP033973
	14	19	100.0	541	2	BF685628	BF685628 602142490
	15	19	100.0	541	4	BG491243	BG491243 602535292
	16	19	100.0	543	2	BF685477	BF685477 602142418
	17	19	100.0	545	4	BG337770	BG337770 602435345
	18	19	100.0	546	4	BG756723	BG756723 602715588
	19	19	100.0	547	2	BF683569	BF683569 602139775
	20	19	100.0	547	4	BG757092	BG757092 602715127
	21	19	100.0	548	4	BF973154	BF973154 602242154
	22	19	100.0	549	2	BF663387	BF663387 602144543
	23	19	100.0	549	4	BG760077	BG760077 602733326
	24	19	100.0	550	4	BF974733	BF974733 602245376
	25	19	100.0	551	2	BF305312	BF305312 601892780
	26	19	100.0	552	4	BG684104	BG684104 602635724
	27	19	100.0	554	4	BF974616	BF974616 602243340
	28	19	100.0	555	4	BG755408	BG755408 602713965
	29	19	100.0	556	4	BG684533	BG684533 602636295
	30	19	100.0	557	2	BF685042	BF685042 602143013
	31	19	100.0	558	4	BG340658	BG340658 602462236
	32	19	100.0	558	4	BG751035	BG751035 602729824
	33	19	100.0	561	2	BF685836	BF685836 602143127
	34	19	100.0	563	4	BM010287	BM010287 603631163
	35	19	100.0	564	2	BF684511	BF684511 602140814
	36	19	100.0	567	4	BG755972	BG755972 602716452
	37	19	100.0	568	4	BG761886	BG761886 602718094
	38	19	100.0	569	4	BG332822	BG332822 602430652
	39	19	100.0	573	4	BG754105	BG754105 602709643
	40	19	100.0	574	4	BG758524	BG758524 602712740
	41	19	100.0	575	4	BG684076	BG684076 602635689
	42	19	100.0	584	4	BG338382	BG338382 602436167
	43	19	100.0	590	4	BG760148	BG760148 602733234
	44	19	100.0	667	4	BG451713	BG451713 NF092F03D
	45	19	100.0	690	4	BG451265	BG451265 NF105B09D

ALIGNMENTS

RESULT 1
AQ215171/c
LOCUS AQ215171 264 bp DNA linear GSS 27-JUL-1999
DEFINITION 1G2 UCBPP-PA14:TnphoA Pseudomonas aeruginosa genomic 5', genomic survey sequence.
ACCESSION AQ215171
VERSION AQ215171.1 GI:4427069
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 264)
AUTHORS Mahajan-Miklos,S.M., Tan,M.-W., Rahme,L.G. and Ausubel,F.M.
TITLE Molecular mechanisms of bacterial virulence elucidated using a Pseudomonas aeruginosa-Caenorhabditis elegans pathogenesis model
JOURNAL Cell 96 (1), 47-56 (1999)
MEDLINE 99142602
PUBMED 9989496
COMMENT Contact: Mahajan-Miklos, S.M.
Department of Molecular Biology
Massachusetts General Hospital
Boston, MA 02114, USA
Fax: 617 726 5950
Email: Mahajan@frodo.mgh.harvard.edu
contains histidine kinse motif
Insert Length: 268 Std Error: 0.00
Seq primer: CGTTACCATGTTAGGAGGTC
Class: transposon-tagged.
FEATURES Location/Qualifiers
source 1. .264
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="UCBPP-PA14 Clinical isolate"
/db_xref="taxon:287"
/clone_lib="UCBPP-PA14:TnphoA"
/note="Vector: pRT731; Transoposon mutagenesis of Pseudomonas aeruginosa UCBPP-PA14 using the transposon TnphoA carried on the suicide plasmid pRT731was performed as previously described (Rahme et al., 1997, Proc. Natl. Acad. Sci. USA, 94: 13245-13250)"
ORIGIN

Query Match 100.0%; Score 19; DB 8; Length 264;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         44 CTGACTCTTATACACAAGT 26


RESULT 2
BE390386
LOCUS       BE390386                482 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  601286635F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3613424 5',
            mRNA sequence.
ACCESSION   BE390386
VERSION     BE390386.1  GI:9335751
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 482)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM274  row: e  column: 09
            High quality sequence start: 11
            High quality sequence stop: 482.
FEATURES             Location/Qualifiers
     source          1. .482
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3613424"
                     /tissue_type="endometrium, adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_44"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies)."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         18 CTGACTCTTATACACAAGT 36


RESULT 3
AQ215170/c
LOCUS       AQ215170                498 bp    DNA     linear   GSS 27-JUL-1999
DEFINITION  8C12 UCBPP-PA14:TnphoA Pseudomonas aeruginosa genomic 5', genomic
            survey sequence.
ACCESSION   AQ215170
VERSION     AQ215170.1  GI:4427068
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Mahajan-Miklos,S.M., Tan,M.-W., Rahme,L.G. and Ausubel,F.M.
  TITLE     Molecular mechanisms of bacterial virulence elucidated using a
            Pseudomonas aeruginosa-Caenorhabditis elegans pathogenesis model
  JOURNAL   Cell 96 (1), 47-56 (1999)
  MEDLINE   99142602
   PUBMED   9989496
COMMENT     Contact: Mahajan-Miklos, S.M.
            Department of Molecular Biology
            Massachusetts General Hospital
            Boston, MA 02114, USA
            Fax: 617 726 5950
            Email: Mahajan@frodo.mgh.harvard.edu
            Insert Length: 500   Std Error: 0.00
            Seq primer: CGTTACCATGTTAGGAGGTC
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="UCBPP-PA14 Clinical isolate"
                     /db_xref="taxon:287"
                     /clone_lib="UCBPP-PA14:TnphoA"
                     /note="Vector: pRT731; Transoposon mutagenesis of
                     Pseudomonas aeruginosa UCBPP-PA14 using the  transposon
                     TnphoA carried on the suicide plasmid pRT731was performed
                     as  previously described (Rahme et al., 1997, Proc. Natl.
                     Acad. Sci.  USA, 94: 13245-13250)"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 8;  Length 498;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         63 CTGACTCTTATACACAAGT 45


RESULT 4
BG338136
LOCUS       BG338136                503 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602435977F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4553947 5',
            mRNA sequence.
ACCESSION   BG338136
VERSION     BG338136.1  GI:13144574
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1252  row: i  column: 20
            High quality sequence start: 18
            High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4553947"
                     /tissue_type="leiomyosarcoma cell line"
```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 503;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 19 CTGACTCTTATACACAAGT 37

RESULT 5
BG340312
LOCUS BG340312 507 bp mRNA linear EST 27-FEB-2001
DEFINITION 602438538F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4556312 5', mRNA sequence.
ACCESSION BG340312
VERSION BG340312.1 GI:13146750
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 507)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1258 row: l column: 09
High quality sequence start: 2
High quality sequence stop: 507.
FEATURES Location/Qualifiers
source 1. .507
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4556312"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 25 CTGACTCTTATACACAAGT 43

RESULT 6
BF685376
LOCUS BF685376 509 bp mRNA linear EST 22-DEC-2000
DEFINITION 602142150F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4303069 5', mRNA sequence.
ACCESSION BF685376
VERSION BF685376.1 GI:11970784
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 509)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1166 row: d column: 14
High quality sequence start: 14
High quality sequence stop: 509.
FEATURES Location/Qualifiers
source 1. .509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4303069"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 25 CTGACTCTTATACACAAGT 43

RESULT 7
BG339203
LOCUS BG339203 509 bp mRNA linear EST 27-FEB-2001
DEFINITION 602437059F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4554998 5', mRNA sequence.
ACCESSION BG339203
VERSION BG339203.1 GI:13145641
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1255  row: e  column: 15
            High quality sequence stop: 509.
FEATURES              Location/Qualifiers
     source           1. .509
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4554998"
                      /tissue_type="leiomyosarcoma cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_46"
                      /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Size-selected >500bp for average insert size
                      1.8kb. Library constructed by Ling Hong in the laboratory
                      of Gerald M. Rubin (University of California, Berkeley)
                      using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                      II RT (Life Technologies). Note: this is a NIH_MGC
                      Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 509;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         25 CTGACTCTTATACACAAGT 43


RESULT 8
BG759369
LOCUS       BG759369                518 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602711887F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4852412 5',
            mRNA sequence.
ACCESSION   BG759369
VERSION     BG759369.1  GI:14070022
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1696  row: m  column: 21
            High quality sequence start: 9
            High quality sequence stop: 518.
FEATURES              Location/Qualifiers
     source           1. .518
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:4852412"
                      /tissue_type="primary B-cells from tonsils (cell line)"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_48"
                      /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                      Site_2: EcoRI; cDNA made by oligo-dT priming.
                      Directionally cloned into EcoRI/XhoI sites using the
                      following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                      for average insert size 1.8kb. Library constructed by Ling
                      Hong in the laboratory of Gerald M. Rubin (University of
                      California, Berkeley) using  ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies).
                      Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         34 CTGACTCTTATACACAAGT 52


RESULT 9
BM010253
LOCUS       BM010253                535 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603631110F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5444691 5',
            mRNA sequence.
ACCESSION   BM010253
VERSION     BM010253.1  GI:16524607
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1924  row: d  column: 04
            High quality sequence start: 22
            High quality sequence stop: 535.
FEATURES              Location/Qualifiers
     source           1. .535
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5444691"
                      /tissue_type="amelanotic melanoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_41"
                      /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Library constructed by Ling Hong in the
                      laboratory of Gerald M. Rubin (University of California,
                      Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
```

Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 535;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 52 CTGACTCTTATACACAAGT 70

RESULT 10
BF971706
LOCUS BF971706 539 bp mRNA linear EST 22-JAN-2001
DEFINITION 602239965F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4328369 5', mRNA sequence.
ACCESSION BF971706
VERSION BF971706.1 GI:12338921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1188 row: b column: 18
High quality sequence start: 31
High quality sequence stop: 539.
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4328369"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 54 CTGACTCTTATACACAAGT 72

RESULT 11
BF974003
LOCUS BF974003 539 bp mRNA linear EST 22-JAN-2001
DEFINITION 602243170F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4331681 5', mRNA sequence.
ACCESSION BF974003
VERSION BF974003.1 GI:12341218
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1196 row: l column: 18
High quality sequence start: 35
High quality sequence stop: 539.
FEATURES Location/Qualifiers
source 1. .539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4331681"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 55 CTGACTCTTATACACAAGT 73

RESULT 12
BG755947
LOCUS BG755947 539 bp mRNA linear EST 15-MAY-2001
DEFINITION 602716423F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4856645 5', mRNA sequence.
ACCESSION BG755947
VERSION BG755947.1 GI:14066600
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

```
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1707  row: n  column: 06
           High quality sequence start: 25
           High quality sequence stop: 539.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4856645"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 539;
  Best Local Similarity  100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         55 CTGACTCTTATACACAAGT 73


RESULT 13
BP033973
LOCUS       BP033973                539 bp    mRNA    linear   EST 19-AUG-2004
DEFINITION  BP033973 Lotus corniculatus var. japonicus flower bud Lotus
            corniculatus var. japonicus cDNA clone MFB001e07_f 3', mRNA
            sequence.
ACCESSION   BP033973
VERSION     BP033973.1  GI:45411133
KEYWORDS    EST.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Characteristics of the Lotus japonicus Gene Repertoire Deduced from
            Large-Scale Expressed Sequence Tag (EST) Analysis
  JOURNAL   Plant Mol. Biol. 54 (3), 405-414 (2004)
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .539
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="mRNA"
                     /isolate="Miyakojima MG-20"
                     /db_xref="taxon:34305"
                     /clone="MFB001e07_f"
                     /tissue_type="flower bud"
                     /clone_lib="Lotus corniculatus var. japonicus flower bud"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 5;  Length 539;
  Best Local Similarity  100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db        115 CTGACTCTTATACACAAGT 133


RESULT 14
BF685628
LOCUS       BF685628                541 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602142490F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4303439 5',
            mRNA sequence.
ACCESSION   BF685628
VERSION     BF685628.1  GI:11971036
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1167  row: c  column: 24
           High quality sequence start: 14
           High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4303439"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 2;  Length 541;
  Best Local Similarity  100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         57 CTGACTCTTATACACAAGT 75


RESULT 15
BG491243
LOCUS       BG491243                541 bp    mRNA    linear   EST 27-MAR-2001
DEFINITION  602535292F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4654126 5',
            mRNA sequence.
ACCESSION   BG491243
```

```
VERSION     BG491243.1  GI:13452755
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1441  row: g  column: 23
            High quality sequence start: 24
            High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4654126"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 541;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         56 CTGACTCTTATACACAAGT 74


RESULT 16
BF685477
LOCUS       BF685477                543 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602142418F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4303421 5',
            mRNA sequence.
ACCESSION   BF685477
VERSION     BF685477.1  GI:11970974
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1167  row: c  column: 06
            High quality sequence start: 28
            High quality sequence stop: 543.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4303421"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         59 CTGACTCTTATACACAAGT 77


RESULT 17
BG337770
LOCUS       BG337770                545 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602435345F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4553194 5',
            mRNA sequence.
ACCESSION   BG337770
VERSION     BG337770.1  GI:13144208
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1250  row: j  column: 11
            High quality sequence start: 34
            High quality sequence stop: 545.
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4553194"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
```

into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 545;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 63 CTGACTCTTATACACAAGT 81

RESULT 18
BG756723
LOCUS BG756723 546 bp mRNA linear EST 15-MAY-2001
DEFINITION 602715588F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855917 5', mRNA sequence.
ACCESSION BG756723
VERSION BG756723.1 GI:14067376
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 546)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1705 row: o column: 22
High quality sequence start: 17
High quality sequence stop: 546.
FEATURES Location/Qualifiers
source 1. .546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4855917"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 62 CTGACTCTTATACACAAGT 80

RESULT 19
BF683569
LOCUS BF683569 547 bp mRNA linear EST 22-DEC-2000
DEFINITION 602139775F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4300819 5', mRNA sequence.
ACCESSION BF683569
VERSION BF683569.1 GI:11968977
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1160 row: f column: 20
High quality sequence start: 30
High quality sequence stop: 547.
FEATURES Location/Qualifiers
source 1. .547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4300819"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 63 CTGACTCTTATACACAAGT 81

RESULT 20
BG757092
LOCUS BG757092 547 bp mRNA linear EST 15-MAY-2001
DEFINITION 602715127F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855279 5', mRNA sequence.
ACCESSION BG757092
VERSION BG757092.1 GI:14067745
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1704  row: e  column: 08
            High quality sequence start: 33
            High quality sequence stop: 547.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4855279"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 547;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         63 CTGACTCTTATACACAAGT 81


RESULT 21
BF973154
LOCUS       BF973154                548 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602242154F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4330884 5',
            mRNA sequence.
ACCESSION   BF973154
VERSION     BF973154.1  GI:12340471
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 548)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1194  row: k  column: 13
            High quality sequence start: 30
            High quality sequence stop: 548.
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4330884"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 548;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         63 CTGACTCTTATACACAAGT 81


RESULT 22
BF663387
LOCUS       BF663387                549 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  602144543F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4297739 5',
            mRNA sequence.
ACCESSION   BF663387
VERSION     BF663387.1  GI:11937282
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1152  row: f  column: 12
            High quality sequence start: 33
            High quality sequence stop: 549.
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4297739"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
```

```
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 549;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         65 CTGACTCTTATACACAAGT 83


RESULT 23
BG760077
LOCUS       BG760077                549 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602733326F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4878655 5',
            mRNA sequence.
ACCESSION   BG760077
VERSION     BG760077.1  GI:14070730
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1765  row: c  column: 08
            High quality sequence start: 34
            High quality sequence stop: 549.
FEATURES             Location/Qualifiers
     source          1. .549
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4878655"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 549;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         65 CTGACTCTTATACACAAGT 83


RESULT 24
BF974733
LOCUS       BF974733                550 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602245376F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4336554 5',
            mRNA sequence.
ACCESSION   BF974733
VERSION     BF974733.1  GI:12341948
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1209  row: g  column: 19
            High quality sequence start: 39
            High quality sequence stop: 550.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4336554"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 550;
  Best Local Similarity   100.0%;  Pred. No. 88;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         66 CTGACTCTTATACACAAGT 84


RESULT 25
BF305312
LOCUS       BF305312                551 bp    mRNA    linear   EST 21-NOV-2000
DEFINITION  601892780F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4138603 5',
            mRNA sequence.
ACCESSION   BF305312
VERSION     BF305312.1  GI:11252194
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1047 row: o column: 20
High quality sequence start: 30
High quality sequence stop: 551.

FEATURES Location/Qualifiers
source 1. .551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4138603"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_17"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 67 CTGACTCTTATACACAAGT 85

RESULT 26
BG684104
LOCUS BG684104 552 bp mRNA linear EST 01-MAY-2001
DEFINITION 602635724F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763741 5', mRNA sequence.
ACCESSION BG684104
VERSION BG684104.1 GI:13915501
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 552)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1619 row: o column: 06
High quality sequence start: 36
High quality sequence stop: 552.

FEATURES Location/Qualifiers
source 1. .552
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4763741"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 552;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 68 CTGACTCTTATACACAAGT 86

RESULT 27
BF974616
LOCUS BF974616 554 bp mRNA linear EST 22-JAN-2001
DEFINITION 602243340F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4334816 5', mRNA sequence.
ACCESSION BF974616
VERSION BF974616.1 GI:12341831
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 554)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM1204 row: o column: 09
High quality sequence start: 38
High quality sequence stop: 554.

FEATURES Location/Qualifiers
source 1. .554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4334816"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 554;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 70 CTGACTCTTATACACAAGT 88

```
RESULT 28
BG755408
LOCUS       BG755408                 555 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602713965F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4854041 5',
            mRNA sequence.
ACCESSION   BG755408
VERSION     BG755408.1  GI:14066061
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 555)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1701  row: a  column: 18
            High quality sequence start: 40
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1. .555
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4854041"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 555;
  Best Local Similarity  100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         71 CTGACTCTTATACACAAGT 89


RESULT 29
BG684533
LOCUS       BG684533                 556 bp    mRNA    linear   EST 01-MAY-2001
DEFINITION  602636295F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4764119 5',
            mRNA sequence.
ACCESSION   BG684533
VERSION     BG684533.1  GI:13915930
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1620  row: n  column: 24
            High quality sequence start: 40
            High quality sequence stop: 556.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4764119"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 556;
  Best Local Similarity  100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         71 CTGACTCTTATACACAAGT 89


RESULT 30
BF685042
LOCUS       BF685042                 557 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602143013F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4303970 5',
            mRNA sequence.
ACCESSION   BF685042
VERSION     BF685042.1  GI:11970450
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1168  row: j  column: 03
            High quality sequence start: 44
            High quality sequence stop: 557.
FEATURES             Location/Qualifiers
     source          1. .557
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4303970"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 557;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         73 CTGACTCTTATACACAAGT 91


RESULT 31
BG340658
LOCUS       BG340658                558 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602462236F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4575033 5',
            mRNA sequence.
ACCESSION   BG340658
VERSION     BG340658.1  GI:13147096
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1285  row: h  column: 10
            High quality sequence start: 41
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4575033"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 558;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         74 CTGACTCTTATACACAAGT 92


RESULT 32
BG751035
LOCUS       BG751035                558 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602729824F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4873589 5',
            mRNA sequence.
ACCESSION   BG751035
VERSION     BG751035.1  GI:14061688
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 558)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1751  row: p  column: 06
            High quality sequence start: 44
            High quality sequence stop: 558.
FEATURES             Location/Qualifiers
     source          1. .558
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4873589"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_43"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 558;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         74 CTGACTCTTATACACAAGT 92


RESULT 33
BF685836
LOCUS       BF685836                561 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602143127F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4304238 5',
            mRNA sequence.
ACCESSION   BF685836
```

```
VERSION     BF685836.1  GI:11971244
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 561)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1169  row: e  column: 07
            High quality sequence start: 46
            High quality sequence stop: 561.
FEATURES             Location/Qualifiers
     source          1. .561
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4304238"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 2;  Length 561;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         78 CTGACTCTTATACACAAGT 96


RESULT 34
BM010287
LOCUS       BM010287                563 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603631163F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5444943 5',
            mRNA sequence.
ACCESSION   BM010287
VERSION     BM010287.1  GI:16524641
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1924  row: n  column: 16
            High quality sequence start: 48
            High quality sequence stop: 563.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5444943"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 19;  DB 4;  Length 563;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         79 CTGACTCTTATACACAAGT 97


RESULT 35
BF684511
LOCUS       BF684511                564 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602140814F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4302075 5',
            mRNA sequence.
ACCESSION   BF684511
VERSION     BF684511.1  GI:11969919
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1163  row: k  column: 04
            High quality sequence start: 52
            High quality sequence stop: 564.
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4302075"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_46"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
```

into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 2; Length 564;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         80 CTGACTCTTATACACAAGT 98
```

RESULT 36
BG755972
LOCUS BG755972 567 bp mRNA linear EST 15-MAY-2001
DEFINITION 602716452F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4856509 5', mRNA sequence.
ACCESSION BG755972
VERSION BG755972.1 GI:14066625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 567)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1707 row: h column: 14
High quality sequence start: 52
High quality sequence stop: 567.

FEATURES Location/Qualifiers
source 1. .567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4856509"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 567;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         83 CTGACTCTTATACACAAGT 101
```

RESULT 37
BG761886
LOCUS BG761886 568 bp mRNA linear EST 15-MAY-2001
DEFINITION 602718094F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841639 5', mRNA sequence.
ACCESSION BG761886
VERSION BG761886.1 GI:14072539
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 568)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1674 row: l column: 24
High quality sequence start: 52
High quality sequence stop: 568.

FEATURES Location/Qualifiers
source 1. .568
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4841639"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"

ORIGIN

Query Match 100.0%; Score 19; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 102
```

RESULT 38
BG332822
LOCUS BG332822 569 bp mRNA linear EST 27-FEB-2001
DEFINITION 602430652F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4548516 5', mRNA sequence.
ACCESSION BG332822
VERSION BG332822.1 GI:13139260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1238  row: g  column: 13
            High quality sequence start: 58
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4548516"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 569;
  Best Local Similarity  100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         85 CTGACTCTTATACACAAGT 103


RESULT 39
BG754105
LOCUS       BG754105                573 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602709643F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4846091 5',
            mRNA sequence.
ACCESSION   BG754105
VERSION     BG754105.1  GI:14064758
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1686  row: f  column: 12
            High quality sequence start: 41
            High quality sequence stop: 573.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4846091"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 4;  Length 573;
  Best Local Similarity  100.0%;  Pred. No. 89;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         89 CTGACTCTTATACACAAGT 107


RESULT 40
BG758524
LOCUS       BG758524                574 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602712740F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853217 5',
            mRNA sequence.
ACCESSION   BG758524
VERSION     BG758524.1  GI:14069177
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1698  row: o  column: 10
            High quality sequence start: 44
            High quality sequence stop: 574.
FEATURES             Location/Qualifiers
     source          1. .574
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4853217"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 100.0%; Score 19; DB 4; Length 574;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         90 CTGACTCTTATACACAAGT 108
```

Search completed: June 13, 2005, 11:36:15
Job time : 1600.5 secs

Katcheves, K.
10/826,953
Seq. Ids 3 & 5

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:53 ; Search time 782.5 Seconds
(without alignments)
1176.548 Million cell updates/sec

Title: US-10-826-573-3
Perfect score: 19
Sequence: 1 ctgactcttatacacaagt 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	100.0	19	6	AR072536	AR072536 Sequence
2	19	100.0	19	6	AR121568	AR121568 Sequence
c 3	19	100.0	19	6	AR159994	AR159994 Sequence
4	19	100.0	19	6	BD251065	BD251065 Method fo
5	19	100.0	19	6	AR214206	AR214206 Sequence
6	19	100.0	19	6	AX080721	AX080721 Sequence
7	19	100.0	19	6	BD064200	BD064200 System fo
c 8	19	100.0	21	6	A51663	A51663 Sequence 7
9	19	100.0	21	6	I01914	I01914 Sequence 3
c 10	19	100.0	21	6	AR183426	AR183426 Sequence
c 11	19	100.0	21	6	AX001597	AX001597 Sequence
c 12	19	100.0	21	6	AX704584	AX704584 Sequence
c 13	19	100.0	30	6	AR363390	AR363390 Sequence
14	19	100.0	38	6	AR364528	AR364528 Sequence
c 15	19	100.0	38	6	AX554970	AX554970 Sequence
16	19	100.0	59	6	I01912	I01912 Sequence 1
c 17	19	100.0	61	12	SYNPIAOR	M60893 Plasmid I-a
c 18	19	100.0	78	12	SYNPIIAOER	M60894 Plasmid II-
19	19	100.0	95	12	SYNIS50RE	M22849 Cloning vec
20	19	100.0	96	6	A29181	A29181 DNA sequenc
c 21	19	100.0	160	6	AX100728	AX100728 Sequence
22	19	100.0	216	12	SYNSLTPHO	M17743 E.coli sltA
c 23	19	100.0	222	6	AX100723	AX100723 Sequence
24	19	100.0	258	6	AX100725	AX100725 Sequence
25	19	100.0	264	6	AR364527	AR364527 Sequence
26	19	100.0	264	12	SYNTN5SUPF	M25496 Synthetic t
c 27	19	100.0	300	6	A51666	A51666 Sequence 10
c 28	19	100.0	300	6	AR183429	AR183429 Sequence
c 29	19	100.0	300	6	AX001600	AX001600 Sequence
c 30	19	100.0	300	6	AX704587	AX704587 Sequence
31	19	100.0	321	6	AX100735	AX100735 Sequence
c 32	19	100.0	321	6	AX100735	AX100735 Sequence
c 33	19	100.0	333	6	AX100722	AX100722 Sequence
c 34	19	100.0	341	6	AX100734	AX100734 Sequence
35	19	100.0	361	6	I44908	I44908 Sequence 56
36	19	100.0	381	6	AX100730	AX100730 Sequence
c 37	19	100.0	395	6	AX100724	AX100724 Sequence
38	19	100.0	397	6	AX100727	AX100727 Sequence
c 39	19	100.0	397	6	AX100732	AX100732 Sequence
c 40	19	100.0	405	6	AX100726	AX100726 Sequence
c 41	19	100.0	431	6	AX100731	AX100731 Sequence
42	19	100.0	490	6	AX100733	AX100733 Sequence
43	19	100.0	520	1	S56312	S56312 imm {dOE, d
44	19	100.0	522	1	PSIS50HUT	Y09435 Pseudomonas
c 45	19	100.0	536	6	AR142541	AR142541 Sequence

ALIGNMENTS

RESULT 1
AR072536
LOCUS AR072536 19 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 7 from patent US 5948622.
ACCESSION AR072536
VERSION AR072536.1 GI:9999300
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Reznikoff,W.S., Goryshin,I.Yu., York,D.L. and Zhou,H.
TITLE System for in vitro transposition
JOURNAL Patent: US 5948622-A 7 07-SEP-1999;
FEATURES Location/Qualifiers
source 1. .19
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 2
AR121568
LOCUS AR121568 19 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6159736.
ACCESSION AR121568
VERSION AR121568.1 GI:14105144
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 19)
AUTHORS Reznikoff,W.S. and Goryshin,I.Y.
TITLE Method for making insertional mutations using a Tn5 synaptic

```
            complex
  JOURNAL   Patent: US 6159736-A 1 12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 3
AR159994/c
LOCUS       AR159994                19 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 13 from patent US 6251655.
ACCESSION   AR159994
VERSION     AR159994.1  GI:16222888
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Minambres Rodriguez,B., Martinez Blanco,H., Rodriguez Olivera,E.,
            Garcia Alonso,B., Fernandez Canon,J.Manuel., Barredo
            Fuente,J.Luis., Diez Garcia,B., Schleissner Sanchez,C., Moreno
            Valle,M.Angel., Salto Maldonado,F. and Luengo Rodriguez,J.Maria.
  TITLE     Process for increasing the production of penicillin G
            (benzylpenicillin) in Penicillium chrysogenum by expression of the
            PCL gene
  JOURNAL   Patent: US 6251655-A 13 26-JUN-2001;
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         19 CTGACTCTTATACACAAGT 1


RESULT 4
BD251065
LOCUS       BD251065                19 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Method for making insertional mutations.
ACCESSION   BD251065
VERSION     BD251065.1  GI:33060835
KEYWORDS    JP 2002531062-A/1.
SOURCE      Transposon Tn5
  ORGANISM  Transposon Tn5
            other sequences; transposons.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Reznikoff,W.S. and Goryshin,I.Y.
  TITLE     Method for making insertional mutations
  JOURNAL   Patent: JP 2002531062-A 1 24-SEP-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT     OS   Transposon Tn5
            PN   JP 2002531062-A/1
            PD   24-SEP-2002
            PF   21-SEP-1999 JP 2000574243
            PR   23-SEP-1998 US    09/159363
            PI   WILLIAM S REZNIKOFF,IGOR Y GORYSHIN
            PC   C12N15/09,C12N9/00,C12N15/01,C12Q1/02//G01N33/15,G01N33/50, PC
              G01N33/566,
            PC   C12N15/00,C12N15/00
            CC   Description of Artificial Sequence: Mosaic
            sequence between OE
            CC           and IE
            CC   sequences
            FH   Key             Location/Qualifiers
            FT   source          1. .19
            FT                   /organism='Transposon Tn5'.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Transposon Tn5"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2411"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 5
AR214206
LOCUS       AR214206                19 bp    DNA     linear   PAT 25-SEP-2002
DEFINITION  Sequence 3 from patent US 6406896.
ACCESSION   AR214206
VERSION     AR214206.1  GI:23311736
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Reznikoff,W.S. and Naumann,T.A.
  TITLE     Transposase enzyme and method for use
  JOURNAL   Patent: US 6406896-A 3 18-JUN-2002;
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 6
AX080721
LOCUS       AX080721                19 bp    DNA     linear   PAT 27-FEB-2001
DEFINITION  Sequence 3 from Patent WO0109363.
ACCESSION   AX080721
VERSION     AX080721.1  GI:13169710
KEYWORDS    .
SOURCE      Transposon Tn5
  ORGANISM  Transposon Tn5
            other sequences; transposons.
REFERENCE   1
  AUTHORS   Reznikoff,W.S. and Naumann,T.A.
  TITLE     Mutant tn5 transposase enzymes and method for their use
  JOURNAL   Patent: WO 0109363-A 3 08-FEB-2001;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Transposon Tn5"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:2411"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 7
BD064200
LOCUS       BD064200                  19 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  System for in vitro transposition using modified TN5 transposase.
ACCESSION   BD064200
VERSION     BD064200.1  GI:22609803
KEYWORDS    JP 2001507565-A/3.
SOURCE      Conus quercinus
  ORGANISM  Conus quercinus
            Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
            Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
            Neogastropoda; Conoidea; Conidae; Conus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Reznikoff,W.S., Goryshin,I.Y. and Zhou,H.
  TITLE     System for in vitro transposition using modified TN5 transposase
  JOURNAL   Patent: JP 2001507565-A 3 12-JUN-2001;
            WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT     PN   JP 2001507565-A/3
            PD   12-JUN-2001
            PF   09-SEP-1997 JP 1998512997
            PR   09-SEP-1996 US    08/814877,02-MAY-1997 US    08/850880 PI
            WILLIAM S REZNIKOFF,IGOR YU GORYSHIN,HONG ZHOU PC
            C12N15/55,C12N9/22,C12N15/90,C12N15/85
            CC   Strandedness: Double;
            CC   Topology: Linear;
            CC   /desc= Tn5 wild outer end
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .19
                        /organism="Conus quercinus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:101313"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 8
A51663/c
LOCUS       A51663                    21 bp    DNA     linear   PAT 10-MAR-1997
DEFINITION  Sequence 7 from Patent WO9617951.
ACCESSION   A51663
VERSION     A51663.1  GI:2304467
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Holden,D.W.
  TITLE     IDENTIFICATION OF GENES
  JOURNAL   Patent: WO 9617951-A 7 13-JUN-1996;
            RPMS TECHNOLOGY LTD (GB)
COMMENT     Other publication AU 4121996 960626.
FEATURES                Location/Qualifiers
     source             1. .21
                        /organism="synthetic construct"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 9
I01914
LOCUS       I01914                    21 bp ss-DNA     linear   PAT 21-MAY-1993
DEFINITION  Sequence 3 from Patent US 4914025.
ACCESSION   I01914
VERSION     I01914.1  GI:271014
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Manoil,C., Beckwith,J., Syvanen,M., Isberg,R.R., Hoffman,C.S. and
            Wright,A.
  TITLE     Export of intra-cellular substances
  JOURNAL   Patent: US 4914025-A 3 03-APR-1990;
            359 Heath St.; Chestnut Hill, MA
FEATURES                Location/Qualifiers
     source             1. .21
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db          2 CTGACTCTTATACACAAGT 20


RESULT 10
AR183426/c
LOCUS       AR183426                  21 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 7 from patent US 6342215.
ACCESSION   AR183426
VERSION     AR183426.1  GI:20227395
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Holden,D.William., Shea,J.Elizabeth. and Hensel,M.
  TITLE     Identification of genes
  JOURNAL   Patent: US 6342215-A 7 29-JAN-2002;
FEATURES                Location/Qualifiers
     source             1. .21
                        /organism="unknown"
                        /mol_type="unassigned DNA"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3
```

```
RESULT 11
AX001597/c
LOCUS       AX001597                  21 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 7 from Patent EP0889120.
ACCESSION   AX001597
VERSION     AX001597.1  GI:7241726
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Holden,D.W.
  TITLE     A micro-organism having reduced adaption to a particular
            environment
  JOURNAL   Patent: EP 0889120-A 7 07-JAN-1999;
            IMP COLLEGE INNOVATIONS LTD (GB)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 12
AX704584/c
LOCUS       AX704584                  21 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 7 from Patent EP1285960.
ACCESSION   AX704584
VERSION     AX704584.1  GI:29538654
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Holden,D.W., Hensel,M. and Shea,J.E.
  TITLE     Virulence genes from Salmonella typhimurium
  JOURNAL   Patent: EP 1285960-A 7 26-FEB-2003;
            Imperial College Innovations Limited (GB) ; Microscience Limited
            (GB)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 13
AR363390/c
LOCUS       AR363390                  30 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 6 from patent US 5212080.
ACCESSION   AR363390
VERSION     AR363390.1  GI:34424886
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Nag,D.K., Huang,H.V. and Berg,D.E.
  TITLE     Method of DNA sequencing using DNA transposon Tn5seql
  JOURNAL   Patent: US 5212080-A 6 18-MAY-1993;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         25 CTGACTCTTATACACAAGT 7


RESULT 14
AR364528
LOCUS       AR364528                  38 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 2 from patent US 5316946.
ACCESSION   AR364528
VERSION     AR364528.1  GI:34427265
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Phadnis,S.H., Huang,H.V. and Berg,D.E.
  TITLE     DNA transposon TN5SUPF in plasmid pBRG1310
  JOURNAL   Patent: US 5316946-A 2 31-MAY-1994;
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 38;
  Best Local Similarity   100.0%;  Pred. No. 94;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 15
AX554970/c
LOCUS       AX554970                  38 bp    DNA     linear   PAT 27-NOV-2002
DEFINITION  Sequence 1 from Patent WO0246444.
ACCESSION   AX554970
VERSION     AX554970.1  GI:25898535
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Goryshin,I.Y., Naumann,T.A. and Reznikoff,W.S.
  TITLE     Double transposition methods for manipulating nucleic acids
  JOURNAL   Patent: WO 0246444-A 1 13-JUN-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="LINKER A (FULL LENGTH)"
ORIGIN
```

```
  Query Match            100.0%;  Score 19;  DB 6;  Length 38;
  Best Local Similarity  100.0%;  Pred. No. 94;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         38 CTGACTCTTATACACAAGT 20


RESULT 16
I01912
LOCUS       I01912                    59 bp ss-DNA     linear   PAT 21-MAY-1993
DEFINITION  Sequence 1 from Patent US 4914025.
ACCESSION   I01912
VERSION     I01912.1  GI:271012
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Manoil,C., Beckwith,J., Syvanen,M., Isberg,R.R., Hoffman,C.S. and
            Wright,A.
  TITLE     Export of intra-cellular substances
  JOURNAL   Patent: US 4914025-A 1 03-APR-1990;
            359 Heath St.; Chestnut Hill, MA
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 59;
  Best Local Similarity  100.0%;  Pred. No. 85;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 17
SYNPIAOR/c
LOCUS       SYNPIAOR                  61 bp    DNA     linear   SYN 27-APR-1993
DEFINITION  Plasmid I-a (containing engineered Tn5 element) O end.
ACCESSION   M60893
VERSION     M60893.1  GI:209014
KEYWORDS    transposon.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Tomcsanyi,T., Berg,C.M., Phadnis,S.H. and Berg,D.E.
  TITLE     Intramolecular transposition by a synthetic IS50 (Tn5) derivative
  JOURNAL   J. Bacteriol. 172 (11), 6348-6354 (1990)
  MEDLINE   91035245
   PUBMED   2172212
COMMENT     Original source text: Synthetic DNA.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     misc_feature    1. .20
                     /note="primer #2"
     misc_feature    37. .55
                     /note="plasmid I-a O end."
ORIGIN

  Query Match            100.0%;  Score 19;  DB 12;  Length 61;
  Best Local Similarity  100.0%;  Pred. No. 85;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         55 CTGACTCTTATACACAAGT 37


RESULT 18
SYNPIIAOER/c
LOCUS       SYNPIIAOER                78 bp    DNA     linear   SYN 27-APR-1993
DEFINITION  Plasmid II-a (containing engineered Tn5 element) O region.
ACCESSION   M60894
VERSION     M60894.1  GI:209026
KEYWORDS    transposon.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Tomcsanyi,T., Berg,C.M., Phadnis,S.H. and Berg,D.E.
  TITLE     Intramolecular transposition by a synthetic IS50 (Tn5) derivative
  JOURNAL   J. Bacteriol. 172 (11), 6348-6354 (1990)
  MEDLINE   91035245
   PUBMED   2172212
COMMENT     Original source text: Synthetic DNA.
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     misc_feature    1. .20
                     /note="primer #3"
     misc_feature    24. .42
                     /note="plasmid II-a O end"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 12;  Length 78;
  Best Local Similarity  100.0%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         72 CTGACTCTTATACACAAGT 54


RESULT 19
SYNIS50RE
LOCUS       SYNIS50RE                 95 bp    DNA     linear   SYN 16-MAR-2000
DEFINITION  Cloning vector DNA, IS50 regulatory element region.
ACCESSION   M22849
VERSION     M22849.1  GI:340772
KEYWORDS    .
SOURCE      Cloning vector pAV10
  ORGANISM  Cloning vector pAV10
            other sequences; artificial sequences; vectors.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Kozlowski,M., Van Brunschot,A., Nash,N. and Davies,R.W.
  TITLE     A novel vector allowing the expression of genes in a wide range of
            gram-negative bacteria
  JOURNAL   Gene 70 (1), 199-204 (1988)
  MEDLINE   89196912
   PUBMED   2853690
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Cloning vector pAV10"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:118301"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 12;  Length 95;
  Best Local Similarity  100.0%;  Pred. No. 77;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19
```

```
RESULT 20
A29181
LOCUS       A29181                    96 bp    DNA     linear   PAT 12-JUL-1995
DEFINITION  DNA sequence  (pTrc99C-phoA) from patent DE3901681.
ACCESSION   A29181
VERSION     A29181.1  GI:1248916
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   .
  JOURNAL   Patent: DE 3901681-A 4 26-JUL-1990;
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 77;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         38 CTGACTCTTATACACAAGT 56


RESULT 21
AX100728/c
LOCUS       AX100728                 160 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 9 from Patent WO0121655.
ACCESSION   AX100728
VERSION     AX100728.1  GI:13619676
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 9 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .160
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 69;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 22
SYNSLTPHO
LOCUS       SYNSLTPHO                216 bp    DNA     linear   SYN 27-APR-1993
DEFINITION  E.coli sltA/phoA fusion gene, partial cds.
ACCESSION   M17743
VERSION     M17743.1  GI:209355
KEYWORDS    fusion protein.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Calderwood,S.B. and Mekalanos,J.J.
  TITLE     Iron regulation of Shiga-like toxin expression in Escherichia coli
            is mediated by the fur locus
  JOURNAL   J. Bacteriol. 169 (10), 4759-4764 (1987)
  MEDLINE   88007425
   PUBMED   3308853
COMMENT     Original source text: E.coli DNA, clone pSC105.
FEATURES             Location/Qualifiers
     source          1. .216
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     CDS             1. .>216
                     /note="sltA/phoA fusion peptide"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA72616.1"
                     /db_xref="GI:209356"
                     /translation="MKIIIFRVLTFFFVIFSVNVVAKEFTLDFSTAKTYVDSLNVIRS
                     AIGTPLPDSYTQVASWTEPFPFCPVLEN"
     sig_peptide     1. .66
                     /note="sltA signal peptide"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 12;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 65;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db        152 CTGACTCTTATACACAAGT 170


RESULT 23
AX100723/c
LOCUS       AX100723                 222 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 4 from Patent WO0121655.
ACCESSION   AX100723
VERSION     AX100723.1  GI:13619671
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 4 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .222
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 222;
  Best Local Similarity   100.0%;  Pred. No. 64;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 24
AX100725
LOCUS       AX100725                 258 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 6 from Patent WO0121655.
ACCESSION   AX100725
VERSION     AX100725.1  GI:13619673
KEYWORDS    .
```

```
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 6 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .258
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db        238 CTGACTCTTATACACAAGT 256


RESULT 25
AR364527
LOCUS       AR364527                264 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 1 from patent US 5316946.
ACCESSION   AR364527
VERSION     AR364527.1  GI:34427264
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 264)
  AUTHORS   Phadnis,S.H., Huang,H.V. and Berg,D.E.
  TITLE     DNA transposon TN5SUPF in plasmid pBRG1310
  JOURNAL   Patent: US 5316946-A 1 31-MAY-1994;
FEATURES             Location/Qualifiers
     source          1. .264
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 264;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 26
SYNTN5SUPF
LOCUS       SYNTN5SUPF              264 bp    DNA     linear   SYN 01-DEC-1994
DEFINITION  Synthetic transposon Tn5supF.
ACCESSION   M25496
VERSION     M25496.1  GI:598401
KEYWORDS    transposon.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 264)
  AUTHORS   Phadnis,S.H., Huang,H.V. and Berg,D.E.
  TITLE     Tn5supF, a 264-base-pair transposon derived from Tn5 for insertion
            mutagenesis and sequencing DNAs cloned in phage lambda
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 86 (15), 5908-5912 (1989)
  MEDLINE   89345574
   PUBMED   2548192
COMMENT     On Dec 7, 1994 this sequence version replaced gi:556436.
            Original source text: Artificial gene DNA.
            Submitted in computer readable form by Phadnis,S.H. 12-JUN-1989.
FEATURES             Location/Qualifiers
     source          1. .264
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
     misc_feature    1. .19
                     /note="sequence required for transposition"
     primer_bind     11. .30
     primer_bind     complement(236. .255)
     misc_feature    246. .264
                     /note="sequence required for transposition"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 12;  Length 264;
  Best Local Similarity   100.0%;  Pred. No. 62;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 27
A51666/c
LOCUS       A51666                  300 bp    DNA     linear   PAT 10-MAR-1997
DEFINITION  Sequence 10 from Patent WO9617951.
ACCESSION   A51666
VERSION     A51666.1  GI:2304470
KEYWORDS    .
SOURCE      Salmonella typhimurium
  ORGANISM  Salmonella typhimurium
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Salmonella.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Holden,D.W.
  TITLE     IDENTIFICATION OF GENES
  JOURNAL   Patent: WO 9617951-A 10 13-JUN-1996;
            RPMS TECHNOLOGY LTD (GB)
COMMENT     Other publication AU 4121996 960626.
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Salmonella typhimurium"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:602"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 60;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 66


RESULT 28
AR183429/c
LOCUS       AR183429                300 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 10 from patent US 6342215.
ACCESSION   AR183429
VERSION     AR183429.1  GI:20227398
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Holden,D.William., Shea,J.Elizabeth. and Hensel,M.
  TITLE     Identification of genes
  JOURNAL   Patent: US 6342215-A 10 29-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="unknown"
```

```
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 60;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 66


RESULT 29
AX001600/c
LOCUS       AX001600                300 bp    DNA     linear   PAT 10-MAR-2000
DEFINITION  Sequence 10 from Patent EP0889120.
ACCESSION   AX001600
VERSION     AX001600.1  GI:7241729
KEYWORDS    .
SOURCE      Salmonella typhimurium
  ORGANISM  Salmonella typhimurium
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Salmonella.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Holden,D.W.
  TITLE     A micro-organism having reduced adaption to a particular
            environment
  JOURNAL   Patent: EP 0889120-A 10 07-JAN-1999;
            IMP COLLEGE INNOVATIONS LTD (GB)
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Salmonella typhimurium"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:602"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 60;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 66


RESULT 30
AX704587/c
LOCUS       AX704587                300 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 10 from Patent EP1285960.
ACCESSION   AX704587
VERSION     AX704587.1  GI:29538657
KEYWORDS    .
SOURCE      Salmonella typhimurium
  ORGANISM  Salmonella typhimurium
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Salmonella.
REFERENCE   1
  AUTHORS   Holden,D.W., Hensel,M. and Shea,J.E.
  TITLE     Virulence genes from Salmonella typhimurium
  JOURNAL   Patent: EP 1285960-A 10 26-FEB-2003;
            Imperial College Innovations Limited (GB) ; Microscience Limited
            (GB)
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Salmonella typhimurium"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:602"
                     /note="partial virulence gene"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 60;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 66


RESULT 31
AX100735
LOCUS       AX100735                321 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 16 from Patent WO0121655.
ACCESSION   AX100735
VERSION     AX100735.1  GI:13619683
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 16 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         69 CTGACTCTTATACACAAGT 87


RESULT 32
AX100735/c
LOCUS       AX100735                321 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 16 from Patent WO0121655.
ACCESSION   AX100735
VERSION     AX100735.1  GI:13619683
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 16 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .321
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match            100.0%;  Score 19;  DB 6;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        199 CTGACTCTTATACACAAGT 181


RESULT 33
AX100722/c
LOCUS       AX100722                333 bp    DNA     linear   PAT 10-APR-2001
```

```
DEFINITION  Sequence 3 from Patent WO0121655.
ACCESSION   AX100722
VERSION     AX100722.1  GI:13619670
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 3 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .333
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 333;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 34
AX100734/c
LOCUS       AX100734                341 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 15 from Patent WO0121655.
ACCESSION   AX100734
VERSION     AX100734.1  GI:13619682
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 15 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .341
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 341;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 35
I44908
LOCUS       I44908                  361 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 56 from patent US 5635617.
ACCESSION   I44908
VERSION     I44908.1  GI:2469621
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 361)
  AUTHORS   Doran,J.L., Kay,W.W., Collinson,S.Karen. and Clouthier,S.C.
  TITLE     Methods and compositions comprising the agfA gene for detection of
            Salmonella
  JOURNAL   Patent: US 5635617-A 56 03-JUN-1997;
FEATURES             Location/Qualifiers
     source          1. .361
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 361;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        335 CTGACTCTTATACACAAGT 353


RESULT 36
AX100730
LOCUS       AX100730                381 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 11 from Patent WO0121655.
ACCESSION   AX100730
VERSION     AX100730.1  GI:13619678
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 11 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .381
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match             100.0%;  Score 19;  DB 6;  Length 381;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        361 CTGACTCTTATACACAAGT 379


RESULT 37
AX100724/c
LOCUS       AX100724                395 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 5 from Patent WO0121655.
ACCESSION   AX100724
VERSION     AX100724.1  GI:13619672
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 5 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .395
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN
```

```
  Query Match           100.0%;  Score 19;  DB 6;  Length 395;
  Best Local Similarity  100.0%;  Pred. No. 57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 38
AX100727
LOCUS       AX100727                397 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 8 from Patent WO0121655.
ACCESSION   AX100727
VERSION     AX100727.1  GI:13619675
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 8 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .397
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 397;
  Best Local Similarity  100.0%;  Pred. No. 57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        377 CTGACTCTTATACACAAGT 395


RESULT 39
AX100732/c
LOCUS       AX100732                397 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 13 from Patent WO0121655.
ACCESSION   AX100732
VERSION     AX100732.1  GI:13619680
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 13 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .397
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 397;
  Best Local Similarity  100.0%;  Pred. No. 57;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 40
AX100726/c
LOCUS       AX100726                405 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 7 from Patent WO0121655.
ACCESSION   AX100726
VERSION     AX100726.1  GI:13619674
KEYWORDS    .
SOURCE      Escherichia coli
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1
  AUTHORS   Tang,C.U.
  TITLE     Virulence gene and protein, and their use
  JOURNAL   Patent: WO 0121655-A 7 29-MAR-2001;
            ISIS INNOVATION LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Escherichia coli"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:562"
ORIGIN

  Query Match           100.0%;  Score 19;  DB 6;  Length 405;
  Best Local Similarity  100.0%;  Pred. No. 56;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


Search completed: June 13, 2005, 10:42:53
Job time : 784.5 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 13, 2005, 09:31:52 ; Search time 200.5 Seconds
(without alignments)
560.973 Million cell updates/sec

Title: US-10-826-573-3
Perfect score: 19
Sequence: 1 ctgactcttatacacaagt 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	19	100.0	19	2	AAT73168	Aat73168 Tn5-deriv
	2	19	100.0	19	2	AAV28399	Aav28399 Transposo
	3	19	100.0	19	2	AAZ06435	Aaz06435 Wildtype
	4	19	100.0	19	3	AAA11739	Aaa11739 Transposo
	5	19	100.0	19	4	AAD21280	Aad21280 Outside e
	6	19	100.0	19	4	AAC85193	Aac85193 Tn5 trans
	7	19	100.0	19	4	AAC91687	Aac91687 Transposo
	8	19	100.0	19	10	AAD58807	Aad58807 Tn5 trans
	9	19	100.0	19	12	ADM95008	Adm95008 Inverted
c	10	19	100.0	19	12	ADM95017	Adm95017 Inverted
	11	19	100.0	19	12	ADQ16516	Adq16516 Transposo
c	12	19	100.0	30	2	AAX01443	Aax01443 O-end-spe
	13	19	100.0	32	2	AAQ65661	Aaq65661 Tn5 trans
	14	19	100.0	32	6	ABK85484	Abk85484 PhoA codi
c	15	19	100.0	38	6	ABK87201	Abk87201 Synthetic
	16	19	100.0	58	6	ABN86339	Abn86339 Modified
	17	19	100.0	60	2	AAQ04280	Aaq04280 Transposo
c	18	19	100.0	60	2	AAQ37195	Aaq37195 BT2 polyl
	19	19	100.0	63	2	AAQ37194	Aaq37194 BT1 polyl
	20	19	100.0	96	2	AAQ05523	Aaq05523 pTrc99C-p
c	21	19	100.0	160	4	AAF79759	Aaf79759 E coli sp
	22	19	100.0	213	2	AAQ40952	Aaq40952 Salmonell
	23	19	100.0	213	2	AAT29054	Aat29054 S. typhim
	24	19	100.0	221	2	AAQ72879	Aaq72879 Salmonell
c	25	19	100.0	222	4	AAF79754	Aaf79754 E coli cs
	26	19	100.0	258	4	AAF79756	Aaf79756 E coli me
	27	19	100.0	264	2	AAQ63801	Aaq63801 Tn5 supF
c	28	19	100.0	300	2	AAT09196	Aat09196 Virulence
	29	19	100.0	321	4	AAF79766	Aaf79766 E coli dg
c	30	19	100.0	321	4	AAF79766	Aaf79766 E coli dg
c	31	19	100.0	333	4	AAF79753	Aaf79753 E coli fi
c	32	19	100.0	341	4	AAF79765	Aaf79765 E coli dg
	33	19	100.0	361	2	AAQ73066	Aaq73066 AgfA sequ
	34	19	100.0	361	2	AAT74141	Aat74141 Salmonell
	35	19	100.0	381	4	AAF79761	Aaf79761 E coli pg
c	36	19	100.0	395	4	AAF79755	Aaf79755 E coli fn
	37	19	100.0	397	4	AAF79758	Aaf79758 E coli fr
c	38	19	100.0	397	4	AAF79763	Aaf79763 E coli em
c	39	19	100.0	405	4	AAF79757	Aaf79757 E coli rp
c	40	19	100.0	431	4	AAF79762	Aaf79762 E coli tr
	41	19	100.0	490	4	AAF79764	Aaf79764 E coli rn
c	42	19	100.0	536	2	AAQ92885	Aaq92885 V. choler
	43	19	100.0	1534	2	AAV28397	Aav28397 Modified
	44	19	100.0	1534	2	AAZ06433	Aaz06433 Modified
	45	19	100.0	1534	2	AAZ22881	Aaz22881 Mutant Tn

ALIGNMENTS

RESULT 1
AAT73168/c
ID AAT73168 standard; DNA; 19 BP.
XX
AC AAT73168;
XX
DT 14-OCT-1998 (first entry)
XX
DE Tn5-derived probe for isolating P. putida pcl gene.
XX
KW Pseudomonas putida U; phenylacetyl-CoA ligase; transposon tagging; Tn5;
KW probe; hybridisation; penicillin G; benzylpenicillin;
KW Penicillium chrysogenum; fungus; ss.
XX
OS Synthetic.
OS Transposon Tn5.
XX
PN WO9735013-A1.
XX
PD 25-SEP-1997.
XX
PF 18-MAR-1997; 97WO-ES000069.
XX
PR 18-MAR-1996; 96ES-00000664.
XX
PA (ANTI) ANTIBIOTICOS SA.
XX
PI Minambres Rodriguez B, Martinez Blanco H, Rodriguez Olivera E;
PI Garcia Alonso B, Fernandez Canon JM, Barredo Fuente JL, Diez Garcia B;
PI Schleissner Sanchez C, Moreno Valle MA, Salto Maldonado F;
PI Luengo Rodriguez JM;
XX
DR WPI; 1997-480218/44.
XX
PT Increasing benzyl-penicillin production by Penicillium chrysogenum - by
PT expressing gene pcl, preferably from Pseudomonas putida U, coding for
PT phenyl:acetyl-CoA ligase.
XX
PS Disclosure; Page 5; 44pp; Spanish.
XX
CC This sequence represents a probe derived from the terminal sequence of
CC the transposon Tn5, which was used to isolate a Tn5 transposon-tagged pcl
CC gene from Pseudomonas putida U. The pcl gene (AAT73167) encodes the

CC enzyme phenylacetyl-CoA ligase (EC 6.2.1.30). The pcl gene was
CC inactivated by transposition mutagenesis and colonies unable to grow on a
CC medium containing a phenylacetyl acid, were screened with the probe. The
CC isolated sequence was then used to screen a genomic library from P.
CC putida U using sequence adjacent to the Tn5 sequence. The pcl gene is
CC used in a method for increasing the production of penicillin G
CC (benzylpenicillin) by Penicillium chrysogenum by expressing the pcl gene
CC in this fungus
XX
SQ Sequence 19 BP; 6 A; 2 C; 5 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 19 CTGACTCTTATACACAAGT 1

RESULT 2
AAV28399
ID AAV28399 standard; DNA; 19 BP.
XX
AC AAV28399;
XX
DT 24-JUL-1998 (first entry)
XX
DE Transposon 5 (Tn5) wild type outside end (OE) sequence.
XX
KW Tn5 transposase; modified; enzyme; in vitro transposition; mutant;
KW target; marker; transposon 5; plasmid pRZTL1; ds.
XX
OS Escherichia coli.
XX
PN WO9810077-A1.
XX
PD 12-MAR-1998.
XX
PF 09-SEP-1997; 97WO-US015941.
XX
PR 09-SEP-1996; 96US-00814877.
PR 02-MAY-1997; 97US-00850880.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Reznikoff WS, Goryshin IY, Zhou H;
XX
DR WPI; 1998-193627/17.
XX
PT Modified Tn5 transposase construct used in novel system for in vitro
PT transposition - used to, e.g. create absolute defective mutants, provide
PT selective markers and to facilitate insertion of specialised DNA
PT sequences into target DNA.
XX
PS Example; Page 21; 73pp; English.
XX
CC This is the transposon 5 (Tn5) wild type outside end (OE) sequence. The
CC invention provides a genetic construct that contains a nucleotide
CC sequence encoding a modified Tn5 transposase enzyme that has both greater
CC avidity for Tn5 OE repeats and is less likely to assume an inactive
CC multimeric form than a wild type Tn5 transposase and a transposable DNA
CC sequence flanked at its 5' and 3' ends by an 18 or 19 base pair flanking
CC DNA sequence comprising nucleotide A at position 10, T at 11 and A at 12.
CC The modified Tn5 transposase and the transposable DNA which is a DNA
CC donor molecule are used in a system for in vitro transposition. The
CC system and method can be used to create absolute defective mutants, to
CC provide selective markers to target DNA, to provide portable regions of
CC homology to a target DNA, to facilitate insertion of specialised DNA
CC sequences into target DNA, to provide primer binding sites or tags for
CC DNA sequencing , to facilitate production of genetic fusion for gene
CC expression studies and protein domain mapping, as well as to bring
CC together other desired combinations of DNA sequences (combinatorial
CC genetics). The modified Tn5 transposase facilitates in vitro
CC transposition reaction rates of at least about 100-fold higher than can
CC be achieved using wild type transposase (as measure in vivo). In vitro
CC transposition using this system can also use donor DNA and target DNA
CC that is circular or linear. The system also requires no outside high
CC energy source and no other protein other than the modified transposase
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 3
AAZ06435
ID AAZ06435 standard; DNA; 19 BP.
XX
AC AAZ06435;
XX
DT 09-NOV-1999 (first entry)
XX
DE Wildtype Outside End (OE) termini.
XX
KW transposase; modified form; wildtype; multimeric; OE termini; IE termini;
KW outside end termini; inside end termini; plasmid; repeat sequence;
KW mutation; ds.
XX
OS Transposon Tn5.
XX
PN US5948622-A.
XX
PD 07-SEP-1999.
XX
PF 06-OCT-1997; 97US-00944916.
XX
PR 09-SEP-1996; 96US-00814877.
PR 02-MAY-1997; 97US-00850880.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Zhou H, York DL, Goryshin IY, Reznikoff WS;
XX
DR WPI; 1999-517947/43.
XX
PT In vitro transposition using a Tn5 based genetic construct.
XX
PS Example 1; Col 16; 48pp; English.
XX
CC Wildtype Outside End (OE, AAZ06435) and Inside End (IE, AAZ06438) were
CC compared and an effort made to randomize the nucleotides at each of the
CC seven positions of difference. A population of oligonucleotides
CC degenerate at each position of difference was created. This resulted in
CC individual oligonucleotides in the population randomly included either
CC the nucleotide of the wildtype OE or the wildtype IE. 128 distinct
CC oligonucleotides were generated, which had the sequence characteristics
CC of both OE and IE and so can be referred to as OE/IE-like sequences. Two
CC of these OE/IE-like sequences are the mutant OE sequences AAZ06436 and
CC AAZ06437
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

```
RESULT 4
AAA11739
ID   AAA11739 standard; DNA; 19 BP.
XX
AC   AAA11739;
XX
DT   21-JUL-2000  (first entry)
XX
DE   Transposon Tn5 outside end DNA fragment.
XX
KW   Transposon; transposase; insertion mutation; synaptic complex; ss.
XX
OS   Transposon Tn5.
XX
PN   WO200017343-A1.
XX
PD   30-MAR-2000.
XX
PF   21-SEP-1999;   99WO-US021960.
XX
PR   23-SEP-1998;   98US-00159363.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Reznikoff WS,  Goryshin IY;
XX
DR   WPI; 2000-283573/24.
XX
PT   Making insertional mutations at random or quasi-random positions in
PT   cellular nucleic acids in target cells, useful for identifying
PT   chromosomal regions involved in expressing or regulating expression of
PT   proteins.
XX
PS   Example; Fig 3; 25pp; English.
XX
CC   This invention describes a novel method (I) for making an insertional
CC   mutation at a random or quasi-random position in cellular nucleic acid in
CC   a target cell. The invention describes a method (II) for forming a
CC   synaptic complex between a Tn5 transposase protein (X) and a
CC   polynucleotide (Y) that comprises a pair of nucleotide sequences adapted
CC   for operably interacting with the Tn5 transposase to form a synaptic
CC   complex and a transposable nucleotide sequence between them, comprising
CC   combining (X) and (Y) in vitro under conditions that disfavor
CC   polynucleotide strand transfer to form the synaptic complex. Methods for
CC   the insertion of exogenous nucleic acids into the nucleic acids of target
CC   cells are used to identify chromosomal regions involved in expressing or
CC   regulating expression of proteins. The same methods may be used in the
CC   development of new therapeutic agents. The transposable polynucleotides
CC   used to form synaptic complexes can consist of transposon apart from any
CC   flanking sequences. This is advantageous in that it reduces the
CC   likelihood of intramolecular transposition and increases the likelihood
CC   of transposition into a target genome. Eliminating donor backbone
CC   sequences from the polynucleotide simplifies preparation of the
CC   transposon sequences to be used in (I). Additionally, the synaptic
CC   complex can form under conditions that disfavor non-productive
CC   intramolecular transposition events. This is advantageous because all of
CC   the synaptic complexes can undergo transposition when combined with
CC   cellular DNA. Little, if any, of the nucleic acid in the synaptic
CC   complexes is inactive. This sequence represents a novel transposon Tn5
CC   outside end DNA fragment described in the method of the invention
XX
SQ   Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 3;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19
```

```
RESULT 5
AAD21280
ID   AAD21280 standard; DNA; 19 BP.
XX
AC   AAD21280;
XX
DT   28-JAN-2002  (first entry)
XX
DE   Outside end (OE) terminal sequence of wild-type Tn5 transposon.
XX
KW   Insertional mutation; synaptic complex; transposon; screening;
KW   outside end; OE; ds.
XX
OS   Unidentified.
XX
PN   US6294385-B1.
XX
PD   25-SEP-2001.
XX
PF   10-AUG-2000; 2000US-00635969.
XX
PR   23-SEP-1998;   98US-00159363.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Goryshin IY,  Reznikoff WS;
XX
DR   WPI; 2001-656171/75.
XX
PT   Making an insertional mutations, especially useful for efficiently
PT   inserting a transposable polynucleotide in a target cell, comprises
PT   introducing into the target cell a synaptic complex.
XX
PS   Disclosure; Fig 3; 11pp; English.
XX
CC   The present invention relates to a method for making an insertional
CC   mutation at a random or quasi-random position in cellular nucleic acid in
CC   a target cell comprising introducing into the target cell a synaptic
CC   complex. The method is particularly useful for efficiently inserting a
CC   transposable polynucleotide at random or quasi-random locations in the
CC   chromosomal or extra-chromosomal nucleic acid of a target cell. The
CC   method may also be used for screening the genome of cells that comprise
CC   an insertional mutation that induces a phenotypic or genotypic change
CC   relative to the cells that are not subject to insertional mutagenesis.
CC   The present sequence is the outside end (OE) terminal sequence of wild-
CC   type Tn5 transposon, used in the exemplification of the invention
XX
SQ   Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19
```

```
RESULT 6
AAC85193
ID   AAC85193 standard; DNA; 19 BP.
XX
AC   AAC85193;
XX
DT   14-MAY-2001  (first entry)
XX
DE   Tn5 transposase outside end (OE) nucleotide sequence.
XX
KW   Tn5 transposon; transposase; mutant; transposition; enzyme; catalyse; ds.
XX
OS   Transposon Tn5.
XX
```

PN WO200109363-A1.
XX
PD 08-FEB-2001.
XX
PF 02-AUG-2000; 2000WO-US021052.
XX
PR 02-AUG-1999; 99US-0146686P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Reznikoff WS, Naumann TA;
XX
DR WPI; 2001-182968/18.
XX
PT Mutant Tn5 transposase enzymes which have preference for Tn5 inside ends
PT over Tn5 outside ends useful for in vitro transpositions and in systems
PT for transposing a transposable DNA sequence in vitro.
XX
PS Disclosure; Fig 1; 37pp; English.
XX
CC The invention relates to a mutant Tn5 transposase protein (I) modified
CC relative to wild-type Tn5 transposase protein, which differs from the
CC wild-type protein at amino acid positions 58 or 372 and which has a
CC preference for Tn5 inside ends (IEs) over Tn5 outside ends (OEs). (I) is
CC useful in a system for transposing a transposable DNA sequence in vitro.
CC The system comprises (I), a donor DNA molecule comprising a transposable
CC DNA sequence which is flanked at its 5' and 3' ends by a wild-type
CC methylated IE sequence and a target DNA molecule into which the
CC transposable element can transpose. It is also useful in a in vitro
CC transposition method which involves combining a donor DNA molecule that
CC comprises a transposable DNA sequence of interest being flanked at its 5'
CC and 3' ends by a wild-type Tn5 IE sequence with a target DNA molecule and
CC (I), in a reaction buffer at a temperature below physiological
CC temperature until the modified transposase binds to the IE sequences and
CC then raising the temperature enzyme to catalyse in vitro transposition.
CC The present sequence represents the nucleotide sequence of Tn5
CC transposase OE
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 7
AAC91687
ID AAC91687 standard; DNA; 19 BP.
XX
AC AAC91687;
XX
DT 27-MAR-2001 (first entry)
XX
DE Transposon Tn5 IS50R O end (3' insertion end).
XX
KW Transposable element; MHC epitope; major histocompatibility complex;
KW intracellular bacterial pathogen; loxP site; Cre recombinase;
KW insertion end; in-frame fusion; detection; antigen;
KW disseminated insertions of class-I epitopes; DICE-I; transposon Tn5;
KW O end; ds.
XX
OS Escherichia coli.
XX
PN WO200071158-A1.
XX
PD 30-NOV-2000.
XX
PF 26-MAY-2000; 2000WO-US014687.
XX
PR 26-MAY-1999; 99US-0136210P.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Heffron FL, Parker DC, Ellefson DD;
XX
DR WPI; 2001-031967/04.
XX
PT Transposable element for detecting an antigenic epitope of a pathogen,
PT comprising 5' and 3' recombining sites, nucleic acid sequences encoding a
PT selectable marker and major histocompatibility complex (MHC) epitope, and
PT an insertion end.
XX
PS Claim 18; Fig 11; 63pp; English.
XX
CC The invention relates to a novel transposable element comprising DNA
CC encoding a selectable marker (e.g., antibiotic resistance) located
CC between a 5' recombining site and a 3' recombining site (e.g., loxP
CC sites); DNA encoding an MHC (major histocompatibility complex) epitope
CC either 5' of the 5' recombining site or 3' of the 3' recombining site;
CC and insertion ends comprising an inverted repeat sequence at the 5' and
CC 3' ends of the transposable element sufficient for integration of the
CC transposable element. The transposable elements of the invention are able
CC to introduce in-frame insertions throughout the chromosome of an
CC intracellular bacterial pathogen. This system "tags" the bacterial gene
CC and resulting protein, allowing the identification of proteins secreted
CC across the membranes of the eukaryotic cell infected by the bacterium. In
CC one embodiment, the transposable elements contain an antibiotic
CC resistance cassette, two minimal loxP recombination sites, an MHC class I
CC or class II epitope, and flanking insertion ends. A transposase, such as
CC the Cre recombinase protein, is expressed in trans from a plasmid, or can
CC be included in the transposable element. The Cre recombinase loops out
CC the intervening sequences containing the antibiotic resistance cassette.
CC When the transposable element inserts within a gene, the resolved
CC insertion places the MHC class I or class II epitope in frame with the
CC gene. The transposable elements of the invention are useful for detecting
CC an antigenic epitope of an intracellular bacterial pathogen, such as
CC Salmonella sp., Mycobacterium tuberculosis and Listeria monocytogenes.
CC Certain embodiments of the technology, termed "disseminated insertions of
CC class-I epitopes" (DICE-I; DICE-II for class II epitopes) allow the rapid
CC and accurate identification of proteins involved in bacterial
CC pathogenesis so that such proteins can be used as vaccine and drug
CC targets. Carrier vaccines may be generated by infecting bacteria with a
CC transposable element of the invention which additionally comprises an
CC antigen associated with a disease, preferably cancer or a viral or
CC bacterial disease, operably linked to the MHC epitope DNA of the
CC transposable element. The present sequence represents a transposon Tn5
CC IS50R O end (3' insertion end) claimed for use in a transposable element
CC of the invention
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 8
AAD58807
ID AAD58807 standard; DNA; 19 BP.
XX
AC AAD58807;
XX
DT 18-DEC-2003 (first entry)
XX
DE Tn5 transposon element DNA #1.
XX
KW Therapeutic protein; gene therapy; transposon; ds.
XX

OS Unidentified.
XX
PN US2003143740-A1.
XX
PD 31-JUL-2003.
XX
PF 15-OCT-2002; 2002US-00272552.
XX
PR 15-OCT-2001; 2001US-0329474P.
PR 08-NOV-2001; 2001US-0344865P.
XX
PA (WOOD/) WOODDELL C.
PA (HERW/) HERWEIJER H.
PA (WOLF/) WOLFF J A.
XX
PI Wooddell C, Herweijer H, Wolff JA;
XX
DR WPI; 2003-645713/61.
XX
PT Integrating nucleic acid into mammalian genome, useful for gene therapy,
PT comprises delivering a complex between nucleic acid containing a
PT transposon and a transposase specific for the transposon.
XX
PS Disclosure; Page 3; 20pp; English.
XX
CC The invention relates to a method of integrating nucleic acid into the
CC genome of mammalian cells. The method involves forming an integrator
CC complex between the nucleic acid containing a transposon and a
CC transposase specific for the transposon and delivering the integrator
CC complex to a mammalian cell. The method and composition is useful for
CC integrating nucleic acid into the genome of mammalian cells, especially
CC nucleic acids encoding therapeutic proteins for gene therapy. The
CC transposon may be used to integrate large DNA molecules, up to 10 kb or
CC larger, into the genome of a mammalian cell. The present sequence is Tn5
CC transposon element DNA (end binding sequence). This sequence is used to
CC illustrate the method of the invention
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 9
ADM95008
ID ADM95008 standard; DNA; 19 BP.
XX
AC ADM95008;
XX
DT 17-JUN-2004 (first entry)
XX
DE Inverted repeat sequence, SEQ ID 3.
XX
KW Inverted repeat sequence; transposable element; transposon Tn5; ds.
XX
OS Synthetic.
XX
PN CA2396611-A1.
XX
PD 31-JAN-2004.
XX
PF 31-JUL-2002; 2002CA-02396611.
XX
PR 31-JUL-2002; 2002CA-02396611.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dyson PJ, Herron P;
XX
DR WPI; 2004-192322/19.
XX
PT New nucleic acid construct comprising inverted repeat sequences of a
PT transposable element and an origin of transfer between the inverted
PT repeat sequences, useful for introducing genetic disruptions in a
PT bacterial genetic material.
XX
PS Claim 4; SEQ ID NO 3; 46pp; English.
XX
CC The present invention relates to a nucleic acid construct (I), which
CC comprises inverted repeat sequences (ADM95006-ADM95009 and ADM95017-
CC ADM95018) of a transposable element and an origin of transfer that lies
CC between the inverted repeat sequences, such that a transposition event
CC involving the inverted repeat sequences will result in the origin of
CC transfer being included in the resultant insertion at the transposition
CC target site. Preferably the inverted repeat sequences are or are derived
CC from the OE and/or IE inverted repeat sequences of the transposon Tn5.
CC The origin of transfer is an oriT, which can be mobilized by the helper
CC plasmids pUZ8002 and pUB307, and has a sequence of ADM95010. The
CC construct comprises a promoterless reporter gene located between the
CC inverted repeat sequences, where the promoterless reporter gene is
CC operatively associated with a ribosome binding site, and the construct
CC further comprises upstream of the reporter gene and ribosome binding site
CC and between the inverted repeat sequences, a translational stop sequence.
CC The construct lacks an origin of replication, is linear, and consists
CC essentially of the inverted repeat sequences and any sequences located
CC between. The nucleic acid construct is useful for introducing genetic
CC disruptions in a bacterial genetic material, particularly that of the
CC Streptomyces species.
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 10
ADM95017/c
ID ADM95017 standard; DNA; 19 BP.
XX
AC ADM95017;
XX
DT 17-JUN-2004 (first entry)
XX
DE Inverted repeat sequence, SEQ ID 12.
XX
KW Inverted repeat sequence; transposable element; transposon Tn5; ds.
XX
OS Synthetic.
XX
PN CA2396611-A1.
XX
PD 31-JAN-2004.
XX
PF 31-JUL-2002; 2002CA-02396611.
XX
PR 31-JUL-2002; 2002CA-02396611.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dyson PJ, Herron P;
XX
DR WPI; 2004-192322/19.
XX
PT New nucleic acid construct comprising inverted repeat sequences of a
PT transposable element and an origin of transfer between the inverted
PT repeat sequences, useful for introducing genetic disruptions in a

PT bacterial genetic material.
XX
PS Claim 4; SEQ ID NO 12; 46pp; English.
XX
CC The present invention relates to a nucleic acid construct (I), which
CC comprises inverted repeat sequences (ADM95006-ADM95009 and ADM95017-
CC ADM95018) of a transposable element and an origin of transfer that lies
CC between the inverted repeat sequences, such that a transposition event
CC involving the inverted repeat sequences will result in the origin of
CC transfer being included in the resultant insertion at the transposition
CC target site. Preferably the inverted repeat sequences are or are derived
CC from the OE and/or IE inverted repeat sequences of the transposon Tn5.
CC The origin of transfer is an oriT, which can be mobilized by the helper
CC plasmids pUZ8002 and pUB307, and has a sequence of ADM95010. The
CC construct comprises a promoterless reporter gene located between the
CC inverted repeat sequences, where the promoterless reporter gene is
CC operatively associated with a ribosome binding site, and the construct
CC further comprises upstream of the reporter gene and ribosome binding site
CC and between the inverted repeat sequences, a translational stop sequence.
CC The construct lacks an origin of replication, is linear, and consists
CC essentially of the inverted repeat sequences and any sequences located
CC between. The nucleic acid construct is useful for introducing genetic
CC disruptions in a bacterial genetic material, particularly that of the
CC Streptomyces species.
XX
SQ Sequence 19 BP; 6 A; 2 C; 5 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 19 CTGACTCTTATACACAAGT 1

RESULT 11
ADQ16516
ID ADQ16516 standard; DNA; 19 BP.
XX
AC ADQ16516;
XX
DT 23-SEP-2004 (first entry)
XX
DE Transposon Tn5 outer element.
XX
KW Transposon Tn5; ss; transposase mediated integration; transposon;
KW transposase; Tn5 outer element; random insertional mutagenesis;
KW RNA polymerase III promoter; U1 snRNA gene.
XX
OS Transposon Tn5.
XX
PN US2004126887-A1.
XX
PD 01-JUL-2004.
XX
PF 08-NOV-2002; 2002US-00291342.
XX
PR 08-NOV-2001; 2001US-0344865P.
XX
PA (WOOD/) WOODDELL C.
PA (HERW/) HERWEIJER H.
PA (WOLF/) WOLFF J A.
XX
PI Wooddell C, Herweijer H, Wolff JA;
XX
DR WPI; 2004-542387/52.
XX
PT Composition useful for enhancing transposase mediated integration of
PT transposon into target nucleic acid, comprising integrator complex, and
PT enhancing reagent.
XX
PS Example; SEQ ID NO 1; 14pp; English.
XX
CC The invention relates to a composition for enhancing transposase mediated
CC integration of a transposon into a target nucleic acid, comprising an
CC integrator complex and an enhancing reagent. The invention also relates
CC to a method of integrating a nucleic acid into a target nucleic acid,
CC involving making a transposon, forming an integrator complex, combining
CC the integrator complex and a cationic enhancing reagent together in
CC solution, and incubating the composition with a target nucleic acid,
CC where the transposase integrates the transposon into the target nucleic
CC acid. The transposase is a hyperactive mutant Tn5 transposase. The Tn5
CC transposase is flanked by elements chosen from Tn5 outer elements, Tn5
CC inner elements and Tn5 mosaic elements. The enhancing reagent is chosen
CC from transfection reagents, polycations, cationic polymers and cationic
CC lipids. The enhancing reagent comprises both cationic proteins and
CC cationic lipids. The composition and the method are useful for providing
CC random insertional mutagenesis, in which integration of a transposon into
CC a target nucleic acid inserts a molecular tag or disrupts a target
CC sequence, where the integration of a molecular tag facilitates cloning,
CC sequencing or identification by providing a detectable marker, and the
CC integration into a coding region disrupts gene function and facilitates
CC study of a gene. The composition is useful for identifying enhancer
CC elements, for sequencing DNA and for integrating large DNA fragments with
CC known ends into a target nucleic acid such as a plasmid, an artificial
CC chromosome or a viral vector. The composition is also useful for
CC integrating e.g. therapeutic genes, siRNA genes, reporter genes, marker
CC or tag sequences, genes containing RNA polymerase III promoters or
CC modified U1 snRNA genes. This sequence represents a transposon Tn5 outer
CC element used in the scope of the invention.
XX
SQ Sequence 19 BP; 6 A; 5 C; 2 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 12; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 12
AAX01443/c
ID AAX01443 standard; DNA; 30 BP.
XX
AC AAX01443;
XX
DT 28-APR-1999 (first entry)
XX
DE O-end-specific primer.
XX
KW Tn5seq1 transposon; RNA transcription; gene hyperexpression;
KW strong promoter; SP6 promoter; T7 promoter; PCR primer; ss.
XX
OS Synthetic.
XX
PN US5869296-A.
XX
PD 09-FEB-1999.
XX
PF 14-JAN-1993; 93US-00004406.
XX
PR 05-OCT-1987; 87US-00105422.
PR 12-APR-1990; 90US-00508382.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Huang HV, Berg DE, Nag DK;
XX
DR WPI; 1999-152772/13.
XX
PT Obtaining hyperexpression of genes in Escherichia coli hosts - by
PT insertion of transposon Tn5seq1 such that the strong SP6 and T7 promoters
PT are adjacent to the host genes.

XX
PS Disclosure; Col 6; 19pp; English.
XX
CC This sequence is a primer used to obtain sequences used in the method of
CC the invention. The method is for RNA transcription, and comprises the
CC insertion of the Tn5seq1 transposon into an E. coli DNA molecule to
CC obtain hyperexpression of genes adjacent to strong promoters SP6 or T7.
CC The transposon is useful for stimulating the transcription of genes
CC adjacent to the heterologous SP6 or T7 promoters in E. coli, for making
CC RNA transcripts in vitro and the hyperexpression or specific
CC transcription of genes (adjacent to the SP6 or T7 ends) in vivo. Tn5seq1
CC offers a less laborious method of sequencing long DNA molecules than
CC current methods such as base-specific chemical cleavage and enzymatic
CC chain termination
XX
SQ Sequence 30 BP; 8 A; 5 C; 8 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 25 CTGACTCTTATACACAAGT 7

RESULT 13
AAQ65661
ID AAQ65661 standard; DNA; 32 BP.
XX
AC AAQ65661;
XX
DT 25-MAR-2003 (revised)
DT 20-JAN-1995 (first entry)
XX
DE Tn5 transposon insertion sequence 50 5' end.
XX
KW 5' terminus; 3' terminus; insertion sequence 50; Tn5 transposon;
KW avirulent; immunogenic; transposon-mediated mutants; turkeys;
KW Pasteurella multocida; pasteurellosis; vaccines; atrophic rhinitis;
KW pneumonia; pigs; enzootic pneumonia; cattle; fowl cholera; ds.
XX
OS Synthetic.
XX
PN WO9411024-A1.
XX
PD 26-MAY-1994.
XX
PF 05-NOV-1993; 93WO-US010600.
XX
PR 06-NOV-1992; 92US-00973070.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Choi KH, Maheswaran SK;
XX
DR WPI; 1994-183160/22.
XX
PT Protecting animals against Pasteurella multocida - by immunising with
PT stable avirulent mutants or with recombinant virulence factor, partic for
PT turkeys.
XX
PS Disclosure; Page 12; 35pp; English.
XX
CC The sequences given in AAQ65661-62 represent the 5' and 3' termini of
CC insertion sequence 50 of the Tn5 transposon. The Tn5 transposon may be
CC used to produce avirulent immunogenic transposon-mediated mutants of
CC Pasteurella multocida. Avirulent mutants produced by this method may be
CC used to protect animals against pasteurellosis in the vaccines of the
CC invention. P. multocida causes atrophic rhinitis and pneumonia in pigs,
CC enzootic pneumonia in cattle and fowl cholera. The vaccines are
CC especially used to treat turkeys. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 32 BP; 9 A; 9 C; 6 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 1 CTGACTCTTATACACAAGT 19

RESULT 14
ABK85484
ID ABK85484 standard; DNA; 32 BP.
XX
AC ABK85484;
XX
DT 21-AUG-2002 (first entry)
XX
DE PhoA coding sequence, PCR primer PhoAfH3.
XX
KW Outer membrane protein; aopB; bacterial cell surface display;
KW microorganism; passenger protein; live vaccine development;
KW library screening; protein purification; bio-contamination;
KW whole cell analysis; PhoA; PCR; primer; ss.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO200236777-A1.
XX
PD 10-MAY-2002.
XX
PF 01-NOV-2001; 2001WO-SG000228.
XX
PR 02-NOV-2000; 2000US-0244902P.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Pan SQ;
XX
DR WPI; 2002-490009/52.
XX
PT Novel Agrobacterium tumefaciens outer membrane polypeptide, termed aopB,
PT useful as carriers to display passenger proteins on surface of bacteria
PT for live vaccine development, library screening, protein purification.
XX
PS Example 6; Page 89; 129pp; English.
XX
CC The present invention relates to the isolation of an Agrobacterium
CC tumefaciens outer membrane protein termed aopB, and the polynucleotide
CC sequence encoding it. The aopB outer membrane protein can be used for
CC bacterial cell surface display of proteins. The invention also provides a
CC method of producing a microorganism on whose surface is displayed a
CC passenger protein. The method is useful in live vaccine development,
CC library screening, protein purification, bio-contamination, and whole
CC cell analysis. The present sequence represents a PCR primer used to
CC amplify a PhoA coding sequence in the examples of the present invention
XX
SQ Sequence 32 BP; 9 A; 7 C; 8 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||||
Db 10 CTGACTCTTATACACAAGT 28

RESULT 15
ABK87201/c

ID ABK87201 standard; DNA; 38 BP.
XX
AC ABK87201;
XX
DT 24-SEP-2002 (first entry)
XX
DE Synthetic full-length transposase-binding linker A.
XX
KW Transposase-interacting inverted repeat sequence pair;
KW transposase enzyme; gene fusion library; transposable element;
KW transposase-binding linker; ds.
XX
OS Synthetic.
XX
PN WO200246444-A2.
XX
PD 13-JUN-2002.
XX
PF 05-DEC-2001; 2001WO-US046311.
XX
PR 05-DEC-2000; 2000US-0251482P.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Goryshin IY, Naumann TA, Reznikoff WS;
XX
DR WPI; 2002-527923/56.
XX
PT Transposable polynucleotide for manipulating nucleic acids to produce
PT gene fusions, comprises two or more transposase-interacting inverted
PT repeat sequence pairs.
XX
PS Disclosure; Fig 1; 53pp; English.
XX
CC The present invention relates to a new polynucleotide comprising distinct
CC first and second transposase-interacting inverted repeat sequence pairs.
CC Each pair has a specificity for binding to and interacting with a
CC distinct transposase enzyme, members of the first sequence pair flanking
CC members of the second sequence pair. The invention is useful for
CC producing a gene fusion library and is also useful for deleting a portion
CC of a chromosome and for cloning a portion of a chromosome of a host cell.
CC The invention is further useful for inserting a preselected
CC polynucleotide sequence insert into a chromosome of a host cell .
CC Transposition occurs without regard to the sequences of the nucleic acid
CC into which the transposable elements transpose. Large libraries having a
CC high level of variability can be produced using the polynucleotide of the
CC invention. The present nucleic acid sequence represents the full-length
CC transposase-binding linker A sequence that is part of a transposase-
CC interacting inverted repeat sequence pair, as described above
XX
SQ Sequence 38 BP; 9 A; 7 C; 8 G; 14 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 38 CTGACTCTTATACACAAGT 20

RESULT 16
ABN86339
ID ABN86339 standard; DNA; 58 BP.
XX
AC ABN86339;
XX
DT 08-OCT-2002 (first entry)
XX
DE Modified mini Tn5 KmNx generating mutagenic primer.
XX
KW Analyte; integrated circuit; light detection system; semiconductor;
KW biosensor; ammonia; lux gene; estrogen; microluminometer; mutagenesis;
KW bioluminescence; Tn5; primer; ss.
XX
OS Synthetic.
XX
PN WO200223168-A2.
XX
PD 21-MAR-2002.
XX
PF 12-SEP-2001; 2001WO-US028464.
XX
PR 12-SEP-2000; 2000US-00660581.
XX
PA (UTBA-) UT-BATTELLE LLC.
XX
PI Simpson ML, Paulus MJ, Sayler GS, Applegate BM, Ripp SA;
XX
DR WPI; 2002-566468/60.
XX
PT Apparatus for detecting target analyte e.g., ammonia, has selectively
PT permeable container affixed to substrate capable of holding luminescent
PT microorganism, and semiconductive layer between substrate and container.
XX
PS Disclosure; Fig 10; 230pp; English.
XX
CC The invention relates to an apparatus for detecting target analyte. The
CC apparatus has integrated circuit (IC) including light detection system,
CC selectively permeable container attached to substrate on IC, layer of
CC semiconducting material between the substrate and container,
CC microorganism within the container, which metabolizes selected analyte to
CC emit light, semiconductive layer between substrate and container, and
CC fluid nutrient reservoir. A biosensor, comprising an IC chip comprising a
CC microorganism that metabolizes ammonia and which harbours a lux gene
CC fused with a heterologous promoter gene stably incorporated into the
CC chromosome of the microorganism where the microorganism is held
CC sufficiently close to a light detection system located on the chip to
CC detect light emitted by a lux gene product expressed in the presence of
CC ammonia is useful for detecting the presence of ammonia. A similar
CC biosensor is useful for detecting an estrogen such as estrone, estradiol,
CC estriol or an esterified estrogen. An integrated microluminometer
CC comprising an IC chip that includes a complementary metal oxide
CC semiconductor (CMOS) photodiode a detector and an n-well/p-substrate
CC junction arranged in an array of junctions across the detector active
CC region is useful for measuring bioluminescence. The present sequence
CC represents a primer used for site-directed mutagenesis to generate a
CC modified mini-Tn5
XX
SQ Sequence 58 BP; 17 A; 12 C; 12 G; 17 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 7 CTGACTCTTATACACAAGT 25

RESULT 17
AAQ04280
ID AAQ04280 standard; DNA; 60 BP.
XX
AC AAQ04280;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 20-SEP-1990 (first entry)
XX
DE Transposon phoA sequence.
XX
KW Transposon; Tn5; alkaline phosphatase; phoA; export DNA; ss.
XX
OS Escherichia coli.
XX

```
FH   Key             Location/Qualifiers
FT   misc_feature    1. .49
FT                   /*tag= a
FT                   /label= transposon 5
FT   misc_difference 30. .30
FT                   /*tag= b
FT                   /mod_base= substitution from A to G
FT   CDS             50. .59
FT                   /*tag= c
FT                   /product= "first nine bases for alkaline phosphatase"
XX
PN   US4914025-A.
XX
PD   03-APR-1990.
XX
PF   05-DEC-1985;   85US-00805486.
XX
PR   05-DEC-1985;   85US-00805486.
XX
PA   (MANO/) MANOIL C.
XX
PI   Manoil C,  Beckwith J,  Syvanen M,  Isbert RR,  Hoffman CS,  Wright A;
XX
DR   WPI; 1990-147416/19.
DR   P-PSDB; AAR04512.
XX
PT   Identification of export DNA sequences in transformed bacteria - using
PT   transposon contg.-structural gene for alkaline phosphatase which requires
PT   export DNA for expression.
XX
PS   Disclosure; Page ?; -pp; English.
XX
CC   Pref. the transposon is Tn5 and the detectable gene product is alkaline
CC   phosphatase. At least one transposon insertion sequence has the codon TGG
CC   in frame with, and upstream from, the gene, removing the dependence on
CC   suppressor mutations. Such substitution gives rise to Tn phoA, while the
CC   native sequence is termed Tn phoA (op). Transformants containing an
CC   export DNA sequence and where transposition has occurred can be screened
CC   for their ability to secrete alkaline phosphatase. (Updated on 25-MAR-
CC   2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 60 BP; 10 A; 18 C; 12 G; 19 T; 0 U; 1 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db          2 CTGACTCTTATACACAAGT 20


RESULT 18
AAQ37195/c
ID   AAQ37195 standard; DNA; 60 BP.
XX
AC   AAQ37195;
XX
DT   25-MAR-2003  (revised)
DT   22-JUN-1993  (first entry)
XX
DE   BT2 polylinker used in construction of transposable element.
XX
KW   Transposon; BT2; transducing particles; bacteria; lytic cycle; detection;
KW   Salmonella; ice nucleation gene; ss.
XX
OS   Synthetic.
XX
PN   US5187061-A.
XX
PD   16-FEB-1993.
XX
PF   05-NOV-1990;   90US-00609331.
XX
PR   04-OCT-1988;   88US-00253160.
PR   05-FEB-1990;   90US-00474282.
XX
PA   (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
PI   Gutterson NI,  Tucker WT,  Wolber PK;
XX
DR   WPI; 1993-075717/09.
XX
PT   Transducing particles prodn. carrying heterologous gene - comprises
PT   introducing DNA sequences into bacterial host contg. prophage which is
PT   induced to a lytic cycle releasing the particles.
XX
PS   Disclosure; Fig 1; 22pp; English.
XX
CC   The essential 19 bp terminal sequence of the transposable element Tn5 was
CC   synthesised chemically as two ca. 60-mer oligonucleotides, BT1 and BT2.
CC   The double stranded linker fragment was prepd. using T4 polynucleotide
CC   kinase, with BT1 5'-3' and BT2 3'-5'. The transposable element is used in
CC   a method for producing transducing particles carrying heterologous genes,
CC   which may be used to detect viable bacteria in biological samples, e.g.
CC   Salmonella which have survived in sterilised food. Detection is specific
CC   due to the specificity of the bacteriophage used. The transducing
CC   particles carry a heterogeneous gene capable of altering the bacterial
CC   phenotype e.g. an ice nucleation gene which allows target bacteria to be
CC   detected at veryy low levels. See also AAQ37194. (Updated on 25-MAR-2003
CC   to correct PF field.)
XX
SQ   Sequence 60 BP; 12 A; 17 C; 18 G; 13 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              ||||||||||||||||||
Db         45 CTGACTCTTATACACAAGT 27


RESULT 19
AAQ37194
ID   AAQ37194 standard; DNA; 63 BP.
XX
AC   AAQ37194;
XX
DT   25-MAR-2003  (revised)
DT   22-JUN-1993  (first entry)
XX
DE   BT1 polylinker used in construction of transposable element.
XX
KW   Transposon; BT2; transducing particles; bacteria; lytic cycle; detection;
KW   Salmonella; ice nucleation gene; ss.
XX
OS   Synthetic.
XX
PN   US5187061-A.
XX
PD   16-FEB-1993.
XX
PF   05-NOV-1990;   90US-00609331.
XX
PR   04-OCT-1988;   88US-00253160.
PR   05-FEB-1990;   90US-00474282.
XX
PA   (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
PI   Gutterson NI,  Tucker WT,  Wolber PK;
XX
DR   WPI; 1993-075717/09.
XX
PT   Transducing particles prodn. carrying heterologous gene - comprises
PT   introducing DNA sequences into bacterial host contg. prophage which is
```

```
PT   induced to a lytic cycle releasing the particles.
XX
PS   Disclosure; Fig 1; 22pp; English.
XX
CC   The essential 19 bp terminal sequence of the transposable element Tn5 was
CC   synthesised chemically as two ca. 60-mer oligonucleotides, BT1 and BT2.
CC   The double stranded linker fragment was prepd. using T4 polynucleotide
CC   kinase, with BT1 5'-3' and BT2 3'-5'. The transposable element is used in
CC   a method for producing transducing particles carrying heterologous genes,
CC   which may be used to detect viable bacteria in biological samples, e.g.
CC   Salmonella which have survived in sterilised food. Detection is specific
CC   due to the specificity of the bacteriophage used. The transducing
CC   particles carry a heterogeneous gene capable of altering the bacterial
CC   phenotype e.g. an ice nucleation gene which allows target bacteria to be
CC   detected at veryy low levels. See also AAQ37195. (Updated on 25-MAR-2003
CC   to correct PF field.)
XX
SQ   Sequence 63 BP; 15 A; 18 C; 16 G; 14 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         23 CTGACTCTTATACACAAGT 41


RESULT 20
AAQ05523
ID   AAQ05523 standard; DNA; 96 BP.
XX
AC   AAQ05523;
XX
DT   25-MAR-2003  (revised)
DT   14-DEC-1990  (first entry)
XX
DE   pTrc99C-phoA.
XX
KW   pTrc99C-phoA; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   RBS             4. .7
FT                   /*tag=  a
FT   CDS             16. .96
FT                   /*tag=  b
FT   misc_RNA        38. .85
FT                   /*tag=  c
FT                   /label= IS50L
XX
PN   DE3901681-A.
XX
PD   26-JUL-1990.
XX
PF   21-JAN-1989;   89DE-03901681.
XX
PR   21-JAN-1989;   89DE-03901681.
XX
PA   (BEHW ) BEHRINGWERKE AG.
XX
PI   Knapp S,  Amann E,  Abel K;
XX
DR   WPI; 1990-232260/31.
DR   P-PSDB; AAR96228.
XX
PT   Signal peptide from Bordetella pertussis - causing secretion of
PT   heterologous proteins in E.coli, and expression vectors for isolating and
PT   testing signal sequences.
XX
PS   Disclosure; Page ?; 18pp; German.
XX
CC   A signal sequence-free phoA gene is present in the vector pTrc99C-phoA.
CC   The vector can be used to test the strength of synthetic signal
CC   sequences. Alkaline phosphatase will only be produced if the sequence is
CC   incorporated in the correct reading frame. See also AAQ05397, AAQ05399-
CC   Q05400 and AAQ05521-Q05522, AAQ05525-Q05526. (Updated on 25-MAR-2003 to
CC   correct PI field.)
XX
SQ   Sequence 96 BP; 21 A; 29 C; 23 G; 23 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 8.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         38 CTGACTCTTATACACAAGT 56


RESULT 21
AAF79759/c
ID   AAF79759 standard; DNA; 160 BP.
XX
AC   AAF79759;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli speA gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 19; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 160 BP; 41 A; 40 C; 44 G; 35 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 8.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 22
```

```
AAQ40952
ID   AAQ40952 standard; DNA; 213 BP.
XX
AC   AAQ40952;
XX
DT   25-MAR-2003  (revised)
DT   29-SEP-1993  (first entry)
XX
DE   Salmonella::phoA fusion joint in pZIP-OUT.
XX
KW   Fusion; Salmonella; phoA; alkaline phosphatase; signal; expression;
KW   outer membrane; export; external surface; vector; bipartite;
KW   tripartite fusion; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   -10_signal      19. .24
FT                   /*tag= a
FT                   /note= "putative Pribnow box"
FT   misc_signal     43. .45
FT                   /*tag= b
FT                   /label= stop_codon
FT   CDS             61. .213
FT                   /*tag= e
FT   misc_signal     61. .63
FT                   /*tag= c
FT                   /note= "putative translation start codon"
FT   misc_signal     94. .96
FT                   /*tag= d
FT                   /note= "putative translation start codon"
FT   misc_RNA        146. .195
FT                   /*tag= f
FT                   /label= IS50L
FT   misc_RNA        196. .213
FT                   /*tag= g
FT                   /label= phoA
FT                   /note= "beginning of phoA"
XX
PN   WO9310246-A1.
XX
PD   27-MAY-1993.
XX
PF   12-NOV-1992;   92WO-US009659.
XX
PR   15-NOV-1991;   91US-00792252.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Niesel DW,  Moncrief JS,  Phillips LH;
XX
DR   WPI; 1993-182560/22.
DR   P-PSDB; AAR37547.
XX
PT   DNA encoding exportation polypeptides - and transformed host cells useful
PT   for prodn. of vaccines and immunogens elicited in response to antigens
PT   expressed on the outer membranes of the host cell.
XX
PS   Disclosure; Fig 2B; 74pp; English.
XX
CC   pZIP-OUT directs the export of fusion polypeptides to the outer membrane
CC   and may also direct a heterologous peptide to the external surface of a
CC   gram-negative host cell. pZIP-OUT is a vector which expresses bipartite
CC   fusion which includes a DNA segment capable of exporting the fusion
CC   product to the external membrane of a gram- negative cell. The other part
CC   of the chimeric gene is a phoA gene segment lacking a signal and
CC   expresssion segments. A variety of DNA segments may be inserted into the
CC   phoA segment at suitable restriction sites to create a tripartite fusion.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 213 BP; 49 A; 50 C; 43 G; 71 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        146 CTGACTCTTATACACAAGT 164


RESULT 23
AAT29054
ID   AAT29054 standard; DNA; 213 BP.
XX
AC   AAT29054;
XX
DT   29-NOV-1996  (first entry)
XX
DE   S. typhimurium surface exportation peptide coding sequence.
XX
KW   Surface exportation protein; S. typhimurium; S. typhimurium; E. coli;
KW   antigen; immune response; cholera; rickettsia; influenza; HIV; ss.
XX
OS   Salmonella typhimurium.
XX
FH   Key             Location/Qualifiers
FT   CDS             61. .213
FT                   /*tag= a
XX
PN   WO9611708-A1.
XX
PD   25-APR-1996.
XX
PF   18-OCT-1995;   95WO-US013333.
XX
PR   18-OCT-1994;   94US-00326772.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Niesel DW,  Moncrief JS,  Phillips LH;
XX
DR   WPI; 1996-221764/22.
DR   P-PSDB; AAR97375.
XX
PT   Membrane expression of heterologous genes e.g. cholera toxin B subunit -
PT   using Salmonella or E.coli and DNA encoding surface export protein, used
PT   for the development of vaccines e.g. to cholera, influenza, HIV, etc.
XX
PS   Claim 2; Page 92; 109pp; English.
XX
CC   This sequence encodes a surface exportation protein derived from S.
CC   typhimurium. This sequence was used in the method of the invention for
CC   inducing antigen-specific antibodies. The method comprises oral
CC   administration of S. typhimurium or E. coli transformed with DNA encoding
CC   the antigen and this surface exportation signal. The peptide encoded by
CC   this sequence can be used to direct surface expression of antigens which
CC   elicit an immune response to cholera, rickettsia, influenza and HIV
XX
SQ   Sequence 213 BP; 49 A; 51 C; 43 G; 70 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        146 CTGACTCTTATACACAAGT 164


RESULT 24
AAQ72879
ID   AAQ72879 standard; DNA; 221 BP.
XX
AC   AAQ72879;
XX
```

DT 12-JUL-1995 (first entry)
XX
DE Salmonella::phoA fusion joint in pZIP-OUT.
XX
KW Plasmid; pZIP-OUT; Salmonella; phoA; fusion joint; outer membrane;
KW S.typhimurium; E.coli; periplasmic space; antigen; cholera toxin;
KW subunit B; vaccine; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT TATA_signal 27. .32
FT /*tag= a
FT misc_signal 69. .71
FT /*tag= b
FT /note= "Putative translation start codon"
FT CDS 102. .221
FT /*tag= c
FT /product= "IS50L/phoA fragment"
XX
PN US5356797-A.
XX
PD 18-OCT-1994.
XX
PF 15-NOV-1991; 91US-00792525.
XX
PR 15-NOV-1991; 91US-00792525.
XX
PA (TEXA) UNIV TEXAS.
XX
PI Phillips LH, Niesel DW, Moncrief JS;
XX
DR WPI; 1994-340329/42.
DR P-PSDB; AAR63628.
XX
PT Recombinant expression of heterologous poly:peptide(s) - using DNA
PT encoding exportation poly:peptide(s) for localisation in E. coli or S.
PT typhimurium cell membranes.
XX
PS Claim 1; Fig 2B; 26pp; English.
XX
CC This sequence represents a portion of the plasmid, pZIP-OUT. This portion
CC covers the Salmonella::phoA fusion joint of pZIP-OUT which may be used to
CC direct the products of large segments of heterologous genes to the outer
CC membrane of S.typhimurium or E.coli or to the external surface of the
CC outer membrane or to an inner membrane/ periplasmic space. This can be
CC used for the production of antigenic proteins, eg. cholera toxin subunit
CC B, for vaccine development
XX
SQ Sequence 221 BP; 52 A; 52 C; 44 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 154 CTGACTCTTATACACAAGT 172

RESULT 25
AAF79754/c
ID AAF79754 standard; DNA; 222 BP.
XX
AC AAF79754;
XX
DT 29-MAY-2001 (first entry)
XX
DE E coli csgE gene fragment.
XX
KW Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS Escherichia coli.
XX
PN WO200121655-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-GB003647.
XX
PR 22-SEP-1999; 99EP-00307495.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Tang C;
XX
DR WPI; 2001-266066/27.
XX
PT Peptides useful as targets for antibacterial therapy, are encoded by
PT virulence genes from enteric bacteria that have a role in colonization
PT during infection.
XX
PS Claim 6; Page 18; 23pp; English.
XX
CC The present invention provides the coding sequences of several E. coli
CC proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC used in the diagnosis and treatment of bacterial infection, and disease
CC in prevention in the form of vaccines against E. coli. The present
CC sequence is one of the aforementioned coding sequences
XX
SQ Sequence 222 BP; 64 A; 48 C; 56 G; 54 T; 0 U; 0 Other;

Query Match 100.0%; Score 19; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 21 CTGACTCTTATACACAAGT 3

RESULT 26
AAF79756
ID AAF79756 standard; DNA; 258 BP.
XX
AC AAF79756;
XX
DT 29-MAY-2001 (first entry)
XX
DE E coli metJ gene fragment.
XX
KW Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS Escherichia coli.
XX
PN WO200121655-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-GB003647.
XX
PR 22-SEP-1999; 99EP-00307495.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Tang C;
XX
DR WPI; 2001-266066/27.
XX
PT Peptides useful as targets for antibacterial therapy, are encoded by
PT virulence genes from enteric bacteria that have a role in colonization

```
PT   during infection.
XX
PS   Claim 6; Page 19; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 258 BP; 72 A; 57 C; 59 G; 70 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 8.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        238 CTGACTCTTATACACAAGT 256


RESULT 27
AAQ63801
ID   AAQ63801 standard; DNA; 264 BP.
XX
AC   AAQ63801;
XX
DT   25-MAR-2003  (revised)
DT   12-JAN-1995  (first entry)
XX
DE   Tn5 supF amber-suppressor transposon nucleotide sequence.
XX
KW   Transposon 5; Tn5; supF; amber-suppressor tRNA; pBRG1310; mutagenesis;
KW   sequencing; phage lambda; IS50; Insertion sequence; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    11. .30
FT                   /*tag= a
FT                   /note= "O end of IS50"
FT   CDS             127. .211
FT                   /*tag= b
FT                   /product= "supF amber-suppressor"
FT   misc_feature    236. .255
FT                   /*tag= c
FT                   /note= "I end of IS50"
XX
PN   US5316946-A.
XX
PD   31-MAY-1994.
XX
PF   22-JAN-1990;   90US-00468450.
XX
PR   05-OCT-1987;   87US-00105422.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Berg DE,  Huang HV,  Phadnis SH;
XX
DR   WPI; 1994-176277/21.
XX
PT   Novel plasmid pBRG1310 contg. Tn5supF transposon - used in mutagenesis
PT   and sequencing DNA(s) cloned in phage lambda.
XX
PS   Claim 1; Fig 1B; 10pp; English.
XX
CC   The inventors have derived a small transposon contained within pBRG1310
CC   which is useful for mutagenesis and sequencing DNAs cloned in phage
CC   lambda. The transposon (Tn5supF) comprises 19bp at each end which
CC   correspond to the O- and I-end segments of IS50 (necessary for
CC   transposition); at least one restriction enzyme site positioned less than
CC   20 nucleotides from each of the terminal sequences; and a supF amber-
CC   suppressor tRNA gene insert. (Updated on 25-MAR-2003 to correct PF
CC   field.)
XX
SQ   Sequence 264 BP; 72 A; 73 C; 60 G; 59 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 264;
  Best Local Similarity   100.0%;  Pred. No. 8.9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db          1 CTGACTCTTATACACAAGT 19


RESULT 28
AAT09196/c
ID   AAT09196 standard; DNA; 300 BP.
XX
AC   AAT09196;
XX
DT   06-JAN-1997  (first entry)
XX
DE   Virulence factor sequence with similarity to Yersinia lcrD gene.
XX
KW   Mutant; adaptation; virulence factor; identification; screening; vaccine;
KW   drugs; infection; treatment; ss.
XX
OS   Salmonella typhimurium.
XX
PN   WO9617951-A2.
XX
PD   13-JUN-1996.
XX
PF   11-DEC-1995;   95WO-GB002875.
XX
PR   09-DEC-1994;   94GB-00024921.
PR   31-JAN-1995;   95GB-00001881.
PR   05-MAY-1995;   95GB-00009239.
XX
PA   (RPMS-) RPMS TECHNOLOGY LTD.
XX
PI   Holden DW;
XX
DR   WPI; 1996-287194/29.
XX
PT   Identifying virulence genes in microorganisms - by introducing mutants
PT   with insertion inactivated genes into environment and retrieval and
PT   analysis of mutants.
XX
PS   Claim 32; Fig 6; 131pp; English.
XX
CC   A method for identifying a microorganism having a reduced adaptation to a
CC   particular environment comprising the steps of: (1) providing a plurality
CC   of microorganisms each of which is independently mutated by the
CC   insertional inactivation of a gene with a nucleic acid comprising a
CC   unique marker sequence so that each mutant contains a different marker
CC   sequence, or clones of the said microorganism; (2) providing individually
CC   a stored sample of each mutant produced by step (1) and providing
CC   individually stored nucleic acid comprising the unique marker sequence
CC   from each individual mutant; (3) introducing a plurality of mutants
CC   produced by step (1) into the said particular environment and allowing
CC   those microorganisms which are able to do so to grow in the said
CC   environment; (4) retrieving microorganisms from the said environment or a
CC   selected part thereof and isolating the nucleic acid from the retrieved
CC   microorganisms; (5) comparing any marker sequences in the nucleic acid
CC   isolated in step (4) to the unique marker sequence of each individual
CC   mutant stored as in step (2); and (6) selecting an individual mutant
CC   which does not contain any of the marker sequences as isolated in step
CC   (4). The products and methods can be used for identifying virulence genes
CC   in microorganisms. The mutant microorganisms can be used in vaccines or
CC   to screen for drugs which reduce virulence or compounds useful for
CC   preventing, ameliorating or treating infections in animals or plants.
```

```
CC   This virulence factor sequence was designated s4C3_1_R
XX
SQ   Sequence 300 BP; 93 A; 57 C; 81 G; 69 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 2;  Length 300;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         84 CTGACTCTTATACACAAGT 66


RESULT 29
AAF79766
ID   AAF79766 standard; DNA; 321 BP.
XX
AC   AAF79766;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli dgcB gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 22; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 321 BP; 81 A; 78 C; 74 G; 82 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         69 CTGACTCTTATACACAAGT 87


RESULT 30
AAF79766/c
ID   AAF79766 standard; DNA; 321 BP.
XX
AC   AAF79766;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli dgcB gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 22; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 321 BP; 81 A; 78 C; 74 G; 82 T; 0 U; 6 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        199 CTGACTCTTATACACAAGT 181


RESULT 31
AAF79753/c
ID   AAF79753 standard; DNA; 333 BP.
XX
AC   AAF79753;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli fimH gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
```

```
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 18; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 333 BP; 88 A; 90 C; 83 G; 72 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 333;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 32
AAF79765/c
ID   AAF79765 standard; DNA; 341 BP.
XX
AC   AAF79765;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli dgcA gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 21; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 341 BP; 101 A; 68 C; 78 G; 94 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 19;  DB 4;  Length 341;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 33
AAQ73066
ID   AAQ73066 standard; DNA; 361 BP.
XX
AC   AAQ73066;
XX
DT   21-OCT-2004  (revised)
DT   27-AUG-2003  (revised)
DT   25-MAR-2003  (revised)
DT   26-JUN-1995  (first entry)
XX
DE   AgfA sequence.
XX
KW   Salmonella; AgfA; vaccine; genetic immunization; ds.
XX
OS   Salmonella enteritidis.
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .359
FT                   /*tag=  a
FT                   /note= "AgfA"
FT   misc_feature    37. .60
FT                   /*tag=  b
FT                   /note= "TAF5 primer (pair with TAF6)"
FT   misc_feature    52. .69
FT                   /*tag=  c
FT                   /note= "TAF3 primer (pair with TAF4)"
FT   misc_feature    103. .129
FT                   /*tag=  d
FT                   /note= "TAF6 primer (pair with TAF5)"
FT   misc_feature    292. .361
FT                   /*tag=  e
FT                   /note= "TAF4 primer (pair with TAF3)"
XX
PN   WO9425598-A2.
XX
PD   10-NOV-1994.
XX
PF   26-APR-1994;   94WO-IB000207.
XX
PR   26-APR-1993;   93US-00054452.
XX
PA   (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.
PA   (KING/) KING J.
XX
PI   Kay WW,  Collinson SK,  Clouthier SC,  Doran JL;
XX
DR   WPI; 1994-358275/44.
DR   P-PSDB; AAR62761.
XX
PT   Eliciting an immune response to Salmonella - using attenuated Salmonella
PT   strains, vector constructs, or compsns. contg. fimbrial type proteins.
XX
PS   Disclosure; Fig 7A; 95pp; English.
XX
CC   The DNA encodes the Salmonella enteritis27655-3b TnphoA mutant strain
CC   agfA gene cloned into pUC19. The DNA and isolated proteins are used in
CC   genetic immunization and vaccine compositions, respectively, to elicit an
CC   immune response to Salmonella in animals (e.g. food producing animals)
CC   and humans. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-
CC   AUG-2003 to correct OS field.)
CC
```

```
CC   Revised record issued on 21-OCT-2004 : Correction to OS line
XX
SQ   Sequence 361 BP; 94 A; 93 C; 94 G; 80 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 2;  Length 361;
  Best Local Similarity  100.0%;  Pred. No. 9.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        335 CTGACTCTTATACACAAGT 353


RESULT 34
AAT74141
ID   AAT74141 standard; DNA; 361 BP.
XX
AC   AAT74141;
XX
DT   25-MAR-2003  (revised)
DT   30-SEP-1997  (first entry)
XX
DE   Salmonella enteritidis 27655-3b TnphoA mutant agfA gene fragment.
XX
KW   Enteropathogenic bacteria; enterobacteria; S.enteritidis; antibody; ds.
XX
OS   Salmonella enteritidis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .360
FT                   /*tag= a
FT                   /label= agfA_gene_fragment
FT   primer_bind     16. .60
FT                   /*tag= b
FT                   /label= Primer_TAF5
FT   primer_bind     52. .69
FT                   /*tag= c
FT                   /label= Primer_TAF3
FT   primer_bind     complement(103. .128)
FT                   /*tag= d
FT                   /label= Primer_TAF6
FT   primer_bind     complement(294. .312)
FT                   /*tag= e
FT                   /label= Primer_TAF4
XX
PN   US5635617-A.
XX
PD   03-JUN-1997.
XX
PF   26-APR-1994;   94US-00233788.
XX
PR   26-APR-1993;   93US-00054452.
XX
PA   (UYVI-) UNIV VICTORIA INNOVATION & DEV CORP.
XX
PI   Collinson SK,  Kay WW,  Doran JL;
XX
DR   WPI; 1997-309886/28.
DR   P-PSDB; AAW23569.
XX
PT   Isolated Salmonella gene agfA - used for diagnosis of Salmonella or
PT   enteropathogenic bacteria of the Enterobacteria family.
XX
PS   Claim 1; Col 107-110; 85pp; English.
XX
CC   The present sequence represents an isolated agfA gene fragment derived
CC   from Salmonella enteritidis 27655-3b TnphoA mutant strain. The nucleic
CC   acid can be used to provide diagnostic assays for Salmonella and/or
CC   enteropathogenic bacteria of the family Enterobacteria. It can also be
CC   used to provide proteins and antibodies which can be used for assays. The
CC   nucleic acid sequence can be used to provide probes or primers which can
CC   specifically hybridise to nucleic acid molecules from greater than 99% of
CC   Salmonella strains that are pathogenic to warm-blooded animals relative
CC   to nucleic acid molecules from virtually all other microbial organisms.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 361 BP; 94 A; 93 C; 94 G; 80 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 2;  Length 361;
  Best Local Similarity  100.0%;  Pred. No. 9.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        335 CTGACTCTTATACACAAGT 353


RESULT 35
AAF79761
ID   AAF79761 standard; DNA; 381 BP.
XX
AC   AAF79761;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli pgi gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 20; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 381 BP; 85 A; 99 C; 106 G; 91 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 19;  DB 4;  Length 381;
  Best Local Similarity  100.0%;  Pred. No. 9.1;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db        361 CTGACTCTTATACACAAGT 379


RESULT 36
AAF79755/c
ID   AAF79755 standard; DNA; 395 BP.
XX
AC   AAF79755;
```

XX
DT 29-MAY-2001 (first entry)
XX
DE E coli fnr gene fragment.
XX
KW Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS Escherichia coli.
XX
PN WO200121655-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-GB003647.
XX
PR 22-SEP-1999; 99EP-00307495.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Tang C;
XX
DR WPI; 2001-266066/27.
XX
PT Peptides useful as targets for antibacterial therapy, are encoded by
PT virulence genes from enteric bacteria that have a role in colonization
PT during infection.
XX
PS Claim 6; Page 18; 23pp; English.
XX
CC The present invention provides the coding sequences of several E. coli
CC proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC used in the diagnosis and treatment of bacterial infection, and disease
CC in prevention in the form of vaccines against E. coli. The present
CC sequence is one of the aforementioned coding sequences
XX
SQ Sequence 395 BP; 97 A; 98 C; 89 G; 108 T; 0 U; 3 Other;

Query Match 100.0%; Score 19; DB 4; Length 395;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 21 CTGACTCTTATACACAAGT 3

RESULT 37
AAF79758
ID AAF79758 standard; DNA; 397 BP.
XX
AC AAF79758;
XX
DT 29-MAY-2001 (first entry)
XX
DE E coli frdA gene fragment.
XX
KW Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS Escherichia coli.
XX
PN WO200121655-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-GB003647.
XX
PR 22-SEP-1999; 99EP-00307495.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Tang C;
XX
DR WPI; 2001-266066/27.
XX
PT Peptides useful as targets for antibacterial therapy, are encoded by
PT virulence genes from enteric bacteria that have a role in colonization
PT during infection.
XX
PS Claim 6; Page 19; 23pp; English.
XX
CC The present invention provides the coding sequences of several E. coli
CC proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC used in the diagnosis and treatment of bacterial infection, and disease
CC in prevention in the form of vaccines against E. coli. The present
CC sequence is one of the aforementioned coding sequences
XX
SQ Sequence 397 BP; 72 A; 110 C; 117 G; 94 T; 0 U; 4 Other;

Query Match 100.0%; Score 19; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGACTCTTATACACAAGT 19
||||||||||||||||||
Db 377 CTGACTCTTATACACAAGT 395

RESULT 38
AAF79763/c
ID AAF79763 standard; DNA; 397 BP.
XX
AC AAF79763;
XX
DT 29-MAY-2001 (first entry)
XX
DE E coli emrB gene fragment.
XX
KW Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS Escherichia coli.
XX
PN WO200121655-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-GB003647.
XX
PR 22-SEP-1999; 99EP-00307495.
XX
PA (ISIS-) ISIS INNOVATION LTD.
XX
PI Tang C;
XX
DR WPI; 2001-266066/27.
XX
PT Peptides useful as targets for antibacterial therapy, are encoded by
PT virulence genes from enteric bacteria that have a role in colonization
PT during infection.
XX
PS Claim 6; Page 21; 23pp; English.
XX
CC The present invention provides the coding sequences of several E. coli
CC proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC used in the diagnosis and treatment of bacterial infection, and disease
CC in prevention in the form of vaccines against E. coli. The present
CC sequence is one of the aforementioned coding sequences
XX

```
SQ   Sequence 397 BP; 87 A; 76 C; 95 G; 139 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 19;  DB 4;  Length 397;
  Best Local Similarity  100.0%;  Pred. No. 9.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 39
AAF79757/c
ID   AAF79757 standard; DNA; 405 BP.
XX
AC   AAF79757;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli rpoN gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 19; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 405 BP; 113 A; 91 C; 105 G; 95 T; 0 U; 1 Other;

  Query Match           100.0%;  Score 19;  DB 4;  Length 405;
  Best Local Similarity  100.0%;  Pred. No. 9.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


RESULT 40
AAF79762/c
ID   AAF79762 standard; DNA; 431 BP.
XX
AC   AAF79762;
XX
DT   29-MAY-2001  (first entry)
XX
DE   E coli treB gene fragment.
XX
KW   Virulence gene; K1 polysaccharide capsule; dgcA; dgcB; dgcC; dgcD; dgcE;
KW   rnr; emrB; treB; pgi; adh; speA; frdA; rpoN; metJ; fnr; csgE; fimH; trsC;
KW   trsE; antibacterial; infection; vaccine; attenuated bacterium; ds.
XX
OS   Escherichia coli.
XX
PN   WO200121655-A2.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-GB003647.
XX
PR   22-SEP-1999;   99EP-00307495.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
XX
PI   Tang C;
XX
DR   WPI; 2001-266066/27.
XX
PT   Peptides useful as targets for antibacterial therapy, are encoded by
PT   virulence genes from enteric bacteria that have a role in colonization
PT   during infection.
XX
PS   Claim 6; Page 20-21; 23pp; English.
XX
CC   The present invention provides the coding sequences of several E. coli
CC   proteins, including dgcA, dgcB, dgcC, dgcD, dgcE, rnr, emrB, treB, pgi,
CC   adh, speA, frdA, rpoN, metJ, fnr, csgE, fimH, trsC and trsE. These can be
CC   used in the diagnosis and treatment of bacterial infection, and disease
CC   in prevention in the form of vaccines against E. coli. The present
CC   sequence is one of the aforementioned coding sequences
XX
SQ   Sequence 431 BP; 98 A; 112 C; 104 G; 116 T; 0 U; 1 Other;

  Query Match           100.0%;  Score 19;  DB 4;  Length 431;
  Best Local Similarity  100.0%;  Pred. No. 9.2;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGACTCTTATACACAAGT 19
              |||||||||||||||||||
Db         21 CTGACTCTTATACACAAGT 3


Search completed: June 13, 2005, 10:16:41
Job time : 202.5 secs
```